AN INVESTIGATION OF POLICE PERFORMANCE UTILIZING MENTAL ABILITY SELECTION SCORES, POLICE ACADEMY TRAINING SCORES, AND SUPERVISORY RATINGS OF THE JOB PERFORMANCE OF PATROL OFFICERS

A THESIS

Presented to

The Faculty of the Division of Graduate Studies

By

Richard L. Feehan

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Psychology

Georgia Institute of Technology

March 1977

AN INVESTIGATION OF POLICE PERFORMANCE UTILIZING MENTAL ABILITY SELECTION SCORES, POLICE ACADEMY TRAINING SCORES, AND SUPERVISORY RATINGS OF THE JOB PERFORMANCE OF PATROL OFFICERS

Approved:

C. Michael York, Ph.D., Chairman

William W. Ronan, Ph.D.

Gerald S. Day, D.B.A.

Date approved by Chairman 3/2/77

ACKNOWLEDGMENTS

There are several individuals to whom I wish to express my sincere thanks for contributions to the completion of this thesis.

Special appreciation is extended to the members of my thesis advisory committee--Dr. C. M. York, Dr. W. W. Ronan, and Dr. G. J. Day--for critically reviewing and editing the thesis. They provided valuable guidance and assistance throughout the project. I am especially grateful to my committee chairman, Dr. York, for generously contributing a considerable amount of time and effort in helping obtain access to the sample and the data. I am grateful to Dr. S. A. Mulaik and Dr. C. V. Riche for their assistance in clarifying some technical aspects of the statistical analyses. Also, appreciation is expressed to Dr. M. I. Appelbaum of the L. L. Thurstone Psychometric Laboratory, University of North Carolina for supplying information regarding the use of the Multivariate Analysis of Variance (MANOVA) program.

Appreciation is accorded Dr. F. J. Landy and Dr. J. L. Farr for permission to use their supervisory rating instrument designed to evaluate patrol performance of police officers. This instrument was constructed as part of research supported by Grant No. NI 71-063-G from the National Institute of Law Enforcement and Criminal Justice, the Law Enforcement

Assistance Administration, and the U. S. Department of Justice.

The Planning and Research Section of the Atlanta Police Department provided full cooperation in the implementation of this study. Special thanks are extended to Major D. M. Edwards and Officer L. T. Lombardy for their support and generous assistance throughout the project. I am grateful to Major R. A. Davis and Sergeant E. M. James of the Atlanta Police Academy for explanations regarding the functioning of the academy, and for descriptions of the training school subject areas. I gratefully acknowledge the cooperation, time, and effort of the police captains, lieutenants, and sergeants who participated in this project.

I am very grateful to Mr. J. K. Lauchner and Mr. E. E. Pritchett of the City of Atlanta Personnel Office for their advice and assistance. I also wish to thank Mr. T. L. Talbert of the Atlanta Regional Commission, and Mr. A. P. Schwartz of the Atlanta Board of Education for their suggestions during the planning stage of this project.

Sincere gratitude is extended to several friends and colleagues who helped me with this thesis. Foremost appreciation goes to Dave Pincus, who provided constant encouragement, support, advice, and assistance. Dave Gentry and Stan Slaton assisted in programming the data for statistical analysis. Thanks are also expressed to the graduate students and faculty of the School of Psychology for their suggestions, support, and encouragement.

Finally, to my parents goes my deepest appreciation for their understanding, patience, encouragement, and support.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

LIST OF TABLES

LIST OF TABLES (CONTINUED)

LIST OF TABLES (CONTINUED)

LIST OF TABLES (CONTINUED)

LIST OF TABLES (CONTINUED)

LIST OF TABLES (CONCLUDED)

SUMMARY

Although increasing emphasis is being placed on improving the effectiveness of law enforcement personnel, relatively little validation research has been conducted on the relationships among actual personnel functions of police agencies. The police profession represents one of the most difficult occupations to evaluate systematically due to the lack of congruent performance standards with which to assess individual police officer effectiveness.

Few studies have employed demographic variables in an attempt to determine the importance of such variables as they moderate police performance. Consequently the importance of such research has until fairly recently been largely ignored. However, the significance of demographic variables as an important source of personnel decision-making is being increasingly recognized by governmental agencies and police agencies. Based upon recent law suits alleging discrimination and unfair labor practices, there is a critical need to address the issue of police performance in a comparative manner employing such demographic characteristics as race, sex, and mental ability.

The research addressed in this study included validation of a mental ability selection instrument against training and job performance; validation of training against job performance; and investigation into the manner in which the

demographic variables of race, sex, and mental ability contribute to performance among subgroups of patrol officers. The sample consisted of 227 patrol officers of the Atlanta Police Department. Five primary analyses were performed on the data, and eleven hypotheses were tested. The instruments used were the Otis Lennon Mental Ability Test, police academy training areas, and behaviorally-anchored supervisory rating scales designed to assess police performance. For three of the analyses, sample subgroups were established based upon race, sex, and mental ability. The fourth analysis involved intercorrelating the 22 variables for the total pooled sample of patrol officers. The fifth analysis tested the relationship between the set of training school scores and the set of supervisory ratings.

Results for the first analysis revealed significant overall race, sex, and mental ability effects for the police academy training variables. The white police recruit group received higher scores for the training school courses than did the black police recruit group. Cultural differences appear to be the principal reason for differences obtained between black and white police recruits in an academic setting. Overall, female police recruits received higher scores in training than did male police recruits. This result can be explained by sexual aptitudinal differences, motivational factors, and achievement factors. The final effect in this analysis was that the higher mental ability recruits received higher scores in training than did lower mental ability

recruits. This finding may be explained by previous research indicating a high positive correlation between intelligence and academic achievement.

Results for the second analysis revealed a significant race effect for the supervisory rating variables. The white police officers received more favorable ratings than did the black police officers. This difference was particularly evident in the male lower mental ability groups. It was concluded that training school performance transfers to some degree to on-the-job performance for the race factor, but not for the sex or mental ability factors. Thus, race was a more valid factor than either sex or mental ability when training indices were validated against supervisory rating criteria. No evidence of racial bias on the part of the supervisors was found in this analysis.

Results for the third analysis revealed no significant effects for the supervisory rating variables adjusted for the training school variables by covariance. The variance from the training school variables was controlled by covariance adjustments in order to obtain some indication whether training school performance did or did not transfer to on-the-job performance for the race effect found in the second analysis. The conclusion from the third analysis was that differences between black and white police officer groups in training school performance, as measured by grades in training courses and subject areas, did contribute to differences between black

and white police officer groups in patrol performance, as measured by supervisory ratings.

Results for the fourth analysis revealed that high intercorrelations existed for the supervisory rating variables, but only moderate intercorrelations existed for the training school variables. The Otis-Lennon variable was moderately related with training school variables, but had no apparent validity for predicting patrol performance. The most crucial finding in this analysis was that none of the training school variables was significantly correlated with supervisory rating variables for the total sample.

To further investigate the relationship between training and job performance, Analysis V was undertaken. Results for the fifth analysis revealed that the set of training school variables was not significantly related to the set of supervisory rating variables for the total pooled sample. Several implications emerged from the findings of Analysis IV and V. First, it may be that performance in training school is somewhat independent of on-the-job performance as a patrol officer. Other personal (e.g., interest, motivation) and situational (e.g., relationships with supervisors, work shift, patrol assignments) factors may influence performance assessment more than the degree of competency exhibited in training school. Second, it may be that training school is sampling different performance dimensions than those on which the supervisory rating instrument is measuring, particularly since training

performance is more related to academic endeavors and technical knowledge, while the supervisory rating scales seem to be tapping trait dimensions.

The general conclusion derived from this study is that demographic variables such as race, sex, and mental ability, seem to be more important at moderating initial recruit training than at moderating actual patrol performance. Training performance does not seem to be a particularly valid predictor of patrol performance. While the mental ability selection test appears to relate moderately to training performance, it seems to be invalid as a predictor of job performance as determined by a behaviorally-based evaluation instrument. A true evaluation of performance in patrol work appears to depend on use of a composite of both objective and subjective criteria.

CHAPTER I

INTRODUCTION

The law enforcement profession is undoubtedly an important occupation in society. The principal functions of the police are: (1) Protection of life and property; (2) Recovery of lost and stolen property; (3) Enforcement of laws and ordinances; (4) Prevention and repression of crime; (5) Detection and apprehension of offenders; (6) Regulation and control of traffic; (7) Preservation of peace and order; and (8) Performance of miscellaneous services to the community (Leonard, 1970; Eldefonso, Coffey, and Grace, 1968). The primary objective of police agencies is the protection of the fundamental rights of all citizens (Spreen, 1971).

A salient goal of modern democratic society is represented by the continual endeavor to create and maintain productive, effective, and efficient law enforcement agencies deemed essential to the orderly existence and satisfactory governance of society (Schubert, 1976; Bozza, 1973; Fleek and Newnam, 1969). Personnel research which addresses the recruitment, selection, training, and performance assessment of individuals for the work role of a police officer should be a continuous activity within law enforcement organizations, especially with the current emphasis on

greater professionalization of the law enforcement establishment. Research directed toward psychological aspects of police work represents an important potential source for upgrading the quality of the law enforcement profession.

Many of the psychological studies of law enforcement personnel which have been published are attempts to scientifically investigate personnel functions and the individuals in the law enforcement role. Several studies have addressed critical areas which should be analyzed systematically in an attempt to improve departmental policies, methods, programs, and personnel, and thus, enhance the profession. A crucial aspect to the improvement of personnel functions in law enforcement agencies involves validation research with selection instruments, training programs, and job performance measures.

The specific variables addressed in the present study are measured proficiency in a formal police training curriculum; supervisory evaluations of patrol officer job performance; and sample attributes of race, sex, and mental ability. While race and sex are employed exclusively as independent variables, mental ability is utilized as an independent and dependent variable contingent on the analysis.

Selected Literature Review

The literature appropriate to the present research has been published in a variety of sources, and has been directed toward numerous audiences. The initial compilation of

literature was extensive and thus required systematic reduction. Since much of the content of the present study is psychological in nature, this selective review focuses on the most closely related studies conducted primarily by behavioral scientists. A composition of the major studies judged most relevant to the present research is provided in Table 1. These studies represent field research conducted in the context of municipal police agencies throughout the nation. The studies have been categorized according to police training proficiency and supervisory assessment of police performance with other aforementioned variables (i.e., race, sex, and mental ability) being subsumed under these categories. The studies were categorized with emphasis given to the variable judged most prominent in the research.

Police Training Proficiency

Several studies have utilized police training proficiency as a major variable. Training data are amenable to use both as a predictor and a criterion measure. The popularity of training proficiency in police research can be traced to its availability; its objectivity; its potential predictive utility; its utility as a criterion measure in the validation of selection instruments and in the construct validation of other measures; and its value in the determination of training needs, objectives, and goals.

DuBois and Watson (1950) conducted one of the earliest studies concerned with validating a battery of predictor

Table 1. Selected Studies Which Have Employed Variables Relevant to the Present Study

STUDY	VARIABLES				
	Training Proficiency	Supervisory Assessment	Mental Ability	Race	Sex
Abbatiello (1969)	X	X	X		
Baehr, Furcon, and Froemel (1968)		X	X	X	
Baehr, et al. (1971)		X	X	X	
Bloch and Anderson (1974)	X	X	X	X	X
Cohen and Chaiken (1972)	X	X	X	X	
DuBois and Watson (1950)	X	X	X		
Furcon (1972)		X	X	X	
Hankey (1968)	X	X	X		
Hess (1973)	X	X	X		
Mills, McDevitt, and Tonkin (1966)	X		X		
Mullineaux (1955)	X	X	X		
Pounion (In Pollack, 1964)	X	X	X		
Spencer and Nichols (1971)		X	X	X	

instruments to training and performance criteria. The sample consisted of 129 male recruits obtained from two successive eight-week classes of the St. Louis Police Academy. The subjects were administered the following battery of tests: a departmental Police Aptitude Test, the Army General Classification Test (AGCT), the Cornell Word Form Test, a Figure Matching Test, the Bennett Mechanical Comprehension Test, the Minnesota Paper Form Board Test, an Object-Aperture Test, the Strong Vocational Interest Blank, the Rosenzweig Picture Frustration Test, and a Handwriting and Composition Test. A Screening Board Rating was also used as a predictor of success as a probationary patrolman. The predictor measures were correlated with four criteria--an achievement test, marksmanship score during academy training, final grade in the police academy, and a supervisory trait rating scale completed after ten weeks on duty.

The AGCT--a measure of mental ability--correlated significantly ($r = .54$ and $.50$) with the final academy grade for Classes I and II respectively. The test correlated .15 with marksmanship score and .10 with supervisory rating, neither correlation coefficient being statistically significant. The multiple correlation coefficients, based upon different combinations of predictors, yielded statistically significant relationships for each criterion. The specially constructed Police Aptitude Test and the AGCT were good predictors of academy performance, while the best predictors of marksmanship

were non-verbal aptitude tests. None of the predictors was significantly correlated with the supervisory rating criterion. Cross-validation with a third class of recruits indicated that the predictor battery as a whole was somewhat more predictive in police training than a single test alone. The conclusion was that the combinations of tests useful in prediction differed considerably among the criteria.

Mullineaux (1955) undertook a study to evaluate the techniques employed to select patrolmen for the Baltimore Police Department. The final study group consisted of 47 male patrol officers. The average education level for the sample was completion of the ninth grade. The selection instruments used in the study were a personal interview and the AGCT. These two measures were validated against training proficiency, as measured by academy grades, and job performance, as measured by supervisory ratings at the end of the first three-month work assignment and six-month work assignment. Police academy grades were composed of numerical ratings on the various academy subjects.

Correlations were found of .56 between AGCT scores and average spelling grades, and .60 between AGCT scores and scores received on report writing. The correlation between AGCT scores and final academy training grades was .66, while a correlation coefficient of .73 was reported between AGCT scores and final examination averages involving academic subjects only. Although it was reported that the sample recruits

did satisfactorily in training and at the end of the first six months on the job, relationships between test scores and job performance, and between training performance and job performance, were not reported. It was concluded that, in general, the techniques used for patrol officer selection were fairly satisfactory.

Pollack (1964) has reported on a 1959 study conducted by Pounion. A specially prepared mental ability written test containing vocabulary, reading comprehension, arithmetic reasoning, and observation sections, was administered to police recruits in the Chicago Police Department. The validity analysis of the test was based on a sample of 221 patrolmen appointed after passing the mental ability test, a physical examination, police academy training, and six months of active patrol duty.

Mental ability test scores correlated .50 with AGCT scores, .35 with police academy grades, but only .08 with supervisory ratings completed at the end of the six-month patrol work period. Possible explanations given for the small coefficient between supervisory ratings and mental ability test scores were the brevity of the patrol duty period and lack of supervisory understanding of the methods and purposes of the ratings. Pollack argues that a more reasonable explanation is that since the sample was a selectively screened group that showed small variability on mental ability test scores, further job performance distinctions would depend on non-

intellective traits.

Mills, McDevitt, and Tonkin (1966) examined the use of situational tests for the police recruit selection program of Cincinnati. As part of a comprehensive screening program, three situational tests were administered in 1964 to a group of 62 male police candidates. Forty-two of these recruits successfully completed police academy training, and were termed the "success" group, while the remaining 20 did not successfully complete police training, and were termed the "failure" group. Scores on the AGCT were used as a reference predictor measure. Two performance measures were derived from police academy records--the first was the final rank of each candidate in his class based on weekly examinations and notebook grades during the training period; the second was the recruit's rank based on scores on the Cincinnati Combat Course, a pistol marksmanship trial.

AGCT scores of the recruits correlated significantly ($r = .60$) with final police academy standing. The Cincinnati Combat Course had a slight positive relationship ($r = .09$) with final police academy standing, and a slight negative relationship ($r = -.05$) with AGCT scores of the recruits. Cross-validation with a second recruit group of candidates substantially confirmed findings on the first group regarding predictive efficiency of the tests.

Hess (1973) attempted to empirically validate police entry tests with police academy training proficiency and job

performance. A random sample of patrolmen on the Cincinnati police force was chosen for the study. Academy performance was assessed with score in training school, while job performance was measured with commendations/years, disciplinary actions/years, first year efficiency, most recent efficiency, and peer score. The predictors for the study included the AGCT, the Minnesota Multiphasic Personality Inventory, first oral interview, and interview with a psychological team.

Multiple correlation analyses were performed for the six predicted variables. The highest multiple correlation coefficient was produced for academy score (.452). Using the guidelines set forth by the Equal Employment Opportunity Commission (EEOC), it was concluded that the AGCT is an unsatisfactory instrument to select patrolmen. The Minnesota Multiphasic Personality Inventory met the minimum statistical level of significance, but it was questionable whether it met the practical significance required by the EEOC. The first oral interview was unsatisfactory as a predictor. The interview with the psychological team was the only predictor that correlated significantly with both academy score and a measure of job performance. The ratios of commendations/years and disciplinary actions/years were unsatisfactory as criterion measures.

Hankey (1968) conducted a study aimed at determining the relationship between personality characteristics and police performance. The hypothesis tested was that successful

police officers would show a syndrome of temperament traits different from unsuccessful police officers. The sample consisted of 801 police officers appointed to the Los Angeles Police Department between 1955 and 1959. The predictors were ten trait and two falsification scores from the Guilford-Zimmerman Temperament Survey; scores from the California Test of Mental Maturity; scores from the Wonderlic Personnel Test; rating on the civil service entrance examination; and age, education, military rank, occupation, and law enforcement experience at the time of appointment. The criteria were three measures of recruit academy performance, a work performance rating, three measures reflecting involvement in censurable behavior, three measures relating to commendatory behavior, four items relating to job tenure, and rating on the oral portion of the sergeants' examination.

For discriminant function analysis, seven groups were established from criterion measures. Each group was divided into success and nonsuccess categories, and discriminant functions were used to investigate differences between success and nonsuccess categories of police officers. The seven groups were Academy Instructor Rating, Academy Final Rating, Work Performance Rating, Number of Disciplinary Days-Off, Sergeants' Oral Examination Rating, Time On and Off the Job, and Nature of Termination. Canonical correlations were determined between sets of predictors and sets of criterion variables.

No difference was found between success and nonsuccess

categories except in the Academy Final Rating Group, in which measures of mental ability were the most important contributors to the discriminant function. A highly significant canonical correlation of .47 was obtained between the set of twenty predictors and a set of ten criteria. However, no evidence was found to support the hypothesis that successful police officers show a different syndrome of personality traits than nonsuccessful police officers.

In 1964 the Chicago Police Department, in cooperation with the Chicago Civil Service Commission, started a research program on police applicants. Two studies which have evolved from this program are those of Abbatiello (1969), dealing principally with training proficiency, and Spencer and Nichols (1971), dealing principally with supervisory assessment of job performance.

Abbatiello (1969) used the civil service examination as a predictor of failure to qualify, as determined by an unfavorable background investigation, and as a predictor of subsequent achievement in the police academy, as measured by training grades and instructor ratings. The final study sample consisted of approximately 430 police officers.

Significant correlation coefficients were obtained between civil service examination scores and academy grades ($r = .35$), and between civil service examination scores and instructor ratings ($r = .25$). A significant correlation coefficient of .62 was obtained between the civil service

examination and the Otis Test of Mental Ability, indicating general similarity of abilities being measured. The civil service examination had predicted scholastic performance in the police academy.

Cohen and Chaiken (1972) conducted a comprehensive investigation employing numerous background variables and police performance measures. Information was obtained about the background and performance of 1,915 officers appointed to the New York City Police Department in 1957, of whom 1,608 were still members of the force in 1968 when most of the data were collected. The background variables were categorized as follows: race, age at appointment, intelligence quotients and civil service examination scores, family descriptors, occupational history, military history, personal history, incidents involving court action, evaluation by a background investigator, early measures of recruit performance, and later experience on and off the job. The performance measures consisted of termination of employment, career advancement, departmental awards and commendations, disciplinary actions, absenteeism, void claims of injury, removal of permission to use firearms, and, for detectives only, arrest activity and supervisory performance evaluations. Additionally, an overall performance index was constructed from the other measures.

The relations between predictor variables and individual performance measures, as well as the relations among the performance measures taken as a group, were determined

from cross-tabulations and simple correlations. These tabulations were obtained separately for the black officers and the total active cohort, which predominantly consisted of white officers. The strength of each background variable as a predictor of later performance was determined by multiple linear regression.

Analysis of the race variable indicated that there were no significant differences found on mental ability scores for black and white officers. For black officers, mental ability was positively related to disciplinary actions and absenteeism, but was not related to career type or awards. Thus, black officers with high intelligence had a greater incidence of departmental misconduct charges than average, including high absenteeism, but they did not have above-average career advancement. For white officers, mental ability was positively related to career advancement and awards. Black officers ranked somewhat lower on the rating by the background investigator than white officers. They also accumulated more departmental disciplinary charges than white officers, but there were no significant differences between the black and white subgroups on the numbers of civilian complaints, allegations of harassment, or criminal charges. However, black officers did not progress through civil service ranks as well as white officers, but they did progress into and through the Detective Division better than whites. With regard to education, the black officers were considerably better educated

than the white officers. Salary levels and termination rates for black and white officers were approximately the same. For white officers, a high civil service score was slightly predictive of good grades in the police academy, but not for black officers.

An officer's recruit training score was the most significant predictor of later performance. Officers who scored high on written examinations based on the material presented in police academy training courses were subsequently better performers than those who scored average or low. They advanced more rapidly through special assignments and civil service promotions, had less departmental misconduct and absenteeism, and had more awards than lower-scoring officers. For black officers, recruit score was significantly related only to later career advancement. For white officers, recruit score was significantly related to the majority of performance measures. Marksmanship was a statistically significant, but not very strong, predictor of career type and later disciplinary actions for the total sample. The relationship between marksmanship and performance variables for the black officer subgroup was similar to the relationship for the total sample, but none of the statistical tests was significant.

An officer's rating while on probation was the second strongest predictor of later performance. Officers rated as "unsatisfactory" on some aspect of performance after nine months on the force tended to have more allegations of

misconduct subsequently, of which more were substantiated, than officers without derogatory ratings. Officers with poor probationary evaluations also tended to be absent more frequently than average. For the black officers, the relationship between probationary evaluation and police performance was almost identical to that for the white officers.

Supervisory Assessment of Police Performance

A large body of literature exists with regard to the traditional supervisory responsibility of employee performance evaluation. However, the literature relevant to the systematic observation and evaluation of police officer effectiveness is relatively limited. The actual performance indices used in assessing police performance have varied from numerous objective criteria (e.g., absenteeism, commendations, disciplinary actions, number of arrests made and citations issued) to the more conventional subjective criteria (e.g., ratings by peers and supervisory officers). Although objective criteria tend to be more valid and reliable indicators of performance, subjective criteria often allow assessment of aspects of performance which objective criteria may not have potential to measure. Perhaps the primary advantage for using subjective measures of performance is that it allows appraisal of constructs, traits, and qualities which aren't normally amenable to expedient objective assessment.

Spencer and Nichols (1971) added to the research undertaken by Abbatiello (1969) with police recruit evaluation in

the Chicago Police Department. The final sample for their study consisted of 268 police officers. A projective design was used such that test scores and background data were obtained at the time of application (1964), and performance measures were obtained four years later for those subjects who remained on the force. The primary predictors were a civil service examination, a biographical data sheet, and a Management Psychologists, Inc. (MPI) personality rating based on a personal history form and a sentence completion form. The criteria consisted of failure to qualify because of an unfavorable evaluation on the background investigation, a departmental performance rating, departmental awards, arrests, complaints, lost time index, sick leave abuser list, safety record, achievements, and undesirable characteristics.

The criteria which were most consistently related to the predictor information were failure to qualify and the departmental performance rating. Patrolmen with high departmental performance ratings tended to have favorable MPI ratings and high scores on the civil service examination. The civil service examination was found to have high correlations with three tests of mental ability (the Otis Test of Mental Ability, the Watson-Glaser Test, and the California Test of Mental Maturity) although the correlation coefficients were not reported. The correlation coefficients interrelating the criteria were generally low to moderate. A multiple correlation coefficient of .20 was obtained using the MPI rating,

the civil service examination, and education level to predict departmental performance rating.

The factor having the highest correlation with both failure to qualify and the departmental performance rating was race. Blacks were more likely to fail the background investigation and they tended to obtain lower departmental performance ratings than did whites. The MPI ratings had a relatively low relationship with race, compared with the civil service examination and other predictors, and had sufficient validity for prediction of failure to qualify and the departmental performance rating. The conclusion was that the MPI ratings were more "culture fair" than usual employment tests.

Baehr, Furcon, and Froemel (1968) conducted an extensive study primarily aimed at the development of valid predictors of patrol officer performance. The sample for the study consisted of between 490 and 540 patrolmen, contingent on predictor variables employed, from an original sample of 2,327 officers who had received performance ratings in the Chicago Police Department. The subjects had volunteered to participate and were selected on the basis of scores on a supervisory paired-comparison rating (i.e., if they had a performance index in the top third or bottom third of officers evaluated) in order to obtain an approximately equal group of highly-rated and lowly-rated patrolmen. Many levels of tenure were represented in the sample, with the lowest level being one year.

An extensive battery of group-administered written tests was given to two groups of officers (Wave I and Wave II) differentiated according to the set of geographical districts patrolled. The two groups were tested five months apart to allow for analysis of the test results for the first group (Wave I) so that the test battery could be refined, if necessary, for use with the second group of subjects (Wave II). Two slightly different versions of the test battery were utilized. The test battery contained predictor variables designed to assess intellectual, motivational, and behavioral components. The criterion variables consisted of eight performance measures--two being subjective (a paired-comparison supervisory rating and a departmental performance rating) and six being objective (tenure, awards, complaints, disciplinary actions, arrests, and absenteeism).

The findings for the study, a segment of which has been reported by Baehr, et al. (1971), showed that the subgroups of black and white patrolmen differed significantly on certain test variables. However, with regard to the overall predictors, the two racial groups were more often similar than they were different.

The racial group differences were significant for all but two criteria (tenure and absenteeism). The black patrolmen made significantly more arrests, had significantly more disciplinary actions taken against them, and had a significantly higher number of complaints sustained. The white

patrolmen had significantly more departmental awards. These differences were suggested to have arisen from a situational context, that is, varying conditions in different districts of assignment may have influenced the findings. White patrolmen were generally rated more favorably on the paired-comparison performance rating and the departmental supervisory rating. The researchers caution that the possibility of supervisor bias in favor of the white patrolman should be a consideration in interpreting the results.

The multiple regression analyses yielded higher validity coefficients for racially separate groups than for racially-mixed groups. Moreover, in predicting the two most significant performance criteria, the best cross-validation coefficients were obtained when the racial groups were treated separately. The multiple correlation coefficients indicated significant relationships between the tests and all eight criterion measures. The multiple correlations obtained across the eight criterion measures showed the coefficient for the white patrolmen increased slightly over the estimate for the total sample, while the coefficient for the black patrolmen increased considerably over the estimate for the total sample. Thus, the primary validation results for the racially-mixed subgroups were lower than those for either racial subgroup treated separately. The conclusion was that equitable and effective selection and placement are possible with multi-racial groups through use of specifically validated tests.

Accurate interpretation of test results depends on consideration given to demographic variables.

A follow-up study to the Baehr, Furcon, and Froemel (1968) study has been summarized by Furcon (1972). The predictive validity of the test battery over time was verified. Significant multiple correlation coefficients ranging from .13 to .27 were obtained by assessing the relationship between 1966 patrol officer test battery performance and actual job performance as much as three years later, as measured by seven of the eight performance criteria used in the study published in 1968. The highest multiple correlation between test battery performance and departmental ratings using a sample of newly-appointed patrolmen in the Chicago Police Department covering a subsequent two and one-half year period was only .07. As the original prediction equations were applied to new officers, the validity coefficients were reduced to be of little practical importance.

Bloch and Anderson (1974) conducted a study primarily to determine performance differences between policemen and policewomen of the Washington, D.C. Police Department. The sample consisted of various numbers of newly-assigned male and female patrol officers depending on the measure utilized. The initial evaluation of the work of new female patrol officers was published as a first report by the Police Foundation in 1973. That report presented findings based on four months of work by 80 new women assigned to patrol duty in the

Washington, D. C. Police Department. The second report (Bloch and Anderson, 1974) deals with the patrol work of 86 new female officers and 86 comparison males over a one-year period, and includes the results of the first report. The performance variables included departmental ratings, an Official's Survey Ratings, the Chief's Survey Ratings, civil service scores, and personnel records.

The rating categories for the departmental rating instrument were bearing and behavior, human relations, learning ability, knowledge and skill, acceptance of responsibility, written expression, oral expression, and performance of duty. The scale was composed of the following three categories: 0-3 meant "below average," 4 - 6 meant "effective and competent," and 7 - 8 meant "excellent." There was no statistically significant difference in overall ratings or in any category of rating of both sexes. Female patrol officers averaged 5.3 while the comparative sample of male patrol officers averaged 5.5 in overall rating. However, if the overall ratings for only those females and comparison males who remained in street assignments in August 1973 are compared, a statistically significant difference emerges. Female officers were rated significantly lower than the male officers (average rating of 5.1 for females, and 5.6 for males). This finding resulted from a transfer to inside assignment of 25 female officers, whose average overall rating was 5.5 (above average for female officers), and a transfer to inside assignment of 9 male officers, whose

average overall rating was 4.9 (below average for male officers).

An Official's Survey, completed anonymously, gave each official in their respective police districts the opportunity to rate patrol officers about whom the official had personal knowledge. The scaling for the survey was 0 - 2 (below average), 3-5 (average), and 6 - 8 (above average). In 1973 captains and lieutenants generally gave the female officers better ratings than they had in 1972, but male officers still received significantly higher ratings from these officials for their ability to handle domestic fights (female officers had an average rating of 3.9; male officers, 4.6). Captains and lieutenants in 1973 rated male and female officers about equal in general competence, ability to handle street violence, and ability to handle upset or injured people. However, when 1973 ratings by captains, lieutenants, and sergeants were averaged together, female officers were rated as less competent than male officers on patrol competence, ability to handle domestic fights, and ability to handle street violence, but were rated equally competent in handling upset of injured people.

A rating of male and female patrol officers on patrol skills was also obtained from the Chief's Survey. The survey was distributed to police district commanders, who usually delegated the task to the patrol sergeant most knowledgeable about the officer. The scale ranged from zero (extremely poor) through 4 (average) to 8 (extremely good). The male officers were rated better than female officers on their general

performance of street patrol (an average of 5.4 for male officers and 4.4 for female officers), protecting a partner from violence (an average of 6.1 for male officers and 4.5 for female officers), handling a public fight (an average of 5.2 for male officers and 3.9 for female officers), and handling disorderly males (an average of 5.1 for male officers and 4.0 for female officers). There were no significant differences between the groups in the areas of handling an automobile accident involving an injury, making a crime report, and handling a disorderly female.

The background variables of race and police academy performance were correlated significantly with an officer's overall departmental rating for both male and female officers. White officers had higher ratings than black officers. The officer's average score in the police academy for all subjects was positively correlated with the overall rating.

Both male and female officers with high civil service test scores tended to be more educated and to have higher police academy scores than others of their sex. White officers had higher civil service test scores than black officers. Newly-assigned females' overall departmental ratings were positively related to education and civil service test scores, and female officers with high civil service test scores received higher ratings on their writing ability at the end of their probationary year. Male officers with high civil service test scores were less likely to resign from the department, and male

officers with better departmental ratings received better ratings on the Chief's Survey on their ability to deal with the public and on their general ability to perform street patrol. Male and female officers with better departmental ratings had higher average ratings on the Official's Survey.

The performance of male and female patrol officers was generally similar. Moreover, males and females were similar in education (average of 12.8 years for females and 12.9 for males), civil service test scores (average of 82 for both groups), number of jobs previously held, and in the ratings they received in pre-employment interviews. Newly-assigned females and comparison males hired at the same time also earned similar average scores at the police academy. The principal differences were the following: (1) Female patrol officers made fewer arrests and gave fewer traffic citations; (2) Female patrol officers were less likely to engage in serious unbecoming conduct; and (3) Female patrol officers were somewhat more likely to be assigned to light duty as the result of injuries, but injuries did not cause them to be absent from work more often than male patrol officers.

The authors concluded that it is appropriate, from a performance viewpoint, to hire women for patrol assignments on the same basis as men. There were no reported incidents which cast serious doubt on the ability of women to perform patrol work satisfactorily. Although one of the principal differences between male and female patrol officers was that women made

fewer arrests and gave fewer traffic citations than men, the authors note that the female officers had less opportunity than male officers in these performance measures. The difference in arrest levels was not reflected in the performance ratings given to the male and female officers. These ratings, reflecting overall departmental judgment of the performance of an officer, indicated about equal overall satisfaction with officers of both sexes.

Overview of the Problem Area

The major deficiencies of reported research with law enforcement personnel may be categorized as follows: (1) overemphasis of personality research; (2) the relative scarcity of validation research; (3) the scarcity of research assessing the importance of demographic variables; and (4) the general scarcity of scientific and systematic performance evaluation systems.

Several studies have been descriptive "armchair" analyses of the characteristics and personality traits of the typical police officer. Generally, an inference is made concerning the desirability and undesirability of certain personality traits in police work, and often the presumption is made that certain personality traits cause a police officer to act in a certain manner on the job. Balch (1972) notes that police personality research has tended to be inconsistent and inconclusive. A critical literature review by Smith (1971) indicated that police personality trait studies have been "only marginally

useful." While such descriptive studies may indicate the average characteristics of police officers, or the desirability and undesirability of specific personality traits, their utility is limited because the information derived from such investigations is commonly not associated empirically or objectively with behavioral measures (Landy and Farr, 1975; Balch, 1972; Smith, 1971).

Landy and Farr (1975) note that the effectiveness of the selection program cannot be improved by merely adding measures of individual characteristics unless the relationship between such characteristics and job performance are known. Validation research, characterized by established relationships between predictive and performance indices, is necessary for ensuring the proper application of selection tests (Baehr, Furcon, and Froemel, 1968). Relatively little experimental research has been conducted on the interrelationships among selection, training, and job performance measures for the police profession. Although numerous studies concerning the selection of police candidates have been reported in the literature, relatively few studies have been undertaken which have dealt with validating police training proficiency against subsequent police performance. There has been a general reluctance among personnel researchers to investigate relationships between training measures and job performance measures, and even a greater reluctance to investigate relationships between selection tests and training measures. There appears to be

some uncertainty among researchers concerning the manner in which to deal with training in validation research. A recent Supreme Court decision rendered June 7, 1976 in the Washington vs. Davis case, a case involving alleged discrimination with recruiting and selection procedures in the Washington, D. C. police force, indicated that selection instruments could appropriately be validated against training proficiency. Thus, for police agencies, training proficiency may be utilized as a predictor of job performance and/or as a criterion index predicted by selection tests.

Although actual job performance is a more important criterion than academy training proficiency, determination of the relationship between training measures and job performance measures may provide some indication of training needs as well as training program effectiveness. While failure to find a substantial relationship between training and job performance measures does not necessarily mean that training is ineffective, such evidence might indicate problem areas which may require further investigation. Learning, as demonstrated by training school proficiency, does not ensure that the police recruit will apply the knowledge, skills, and abilities on the job. Having obtained the necessary ability to adequately perform the job of a police officer, subsequent job performance may reflect other characteristics which are neither incorporated in the training program nor reflected in training performance. Job performance may reflect the impact of various personal

and environmental factors (e.g., attitude, motivation, frustration, stress, disillusionment, job satisfaction, supervision, quality of field training, nature of work assignment) which tend to vary widely for individual police officers. Validation research, using training as a predictor and job performance measures as a criterion variable, may help determine personal problem areas as well as organizational problems.

The major impetus for the investigation of demographic variables (e.g., race and sex) in personnel research has been provided by the government. Comparative investigations of research samples using demographic variables represents a critical area for differential validation research. Research utilizing demographic variables may help in the formulation of recruitment strategies; may aid in the determination of selection program modifications; may suggest whether supplemental training is needed; may provide moderators of performance which should be considered in performance system formulation; may suggest job classifications in which particular individuals would be more productive and/or more satisfied; may provide evidence of discriminatory treatment; and may suggest bases on which to make personnel management decisions.

The investigation of demographic variables has become a major issue in validation research. The Department of Labor (1968) issued a federal executive order which requires all organizations receiving federal funding to provide evidence of the validity of employment tests, and to take affirmative

action to ensure that such tests are not used to discriminate against minority groups. The EEOC guidelines prohibit the use of any test that adversely affects the hiring of classes of persons protected by the Civil Rights Act unless the test has been validated and exemplifies a high degree of utility, and suitable alternative tests are unavailable (Federal Register, EEOC Guidelines on Employee Selection Procedures, 1970).

In addition to governmental agencies, private organizations and interest groups are exerting coercive pressures on personnel agencies to show the job-relatedness of screening devices, and their relative usefulness in separating the potentially satisfactory worker from the potentially unsatisfactory worker. In situations involving disproportionate hiring, the equitability and impartiality of a screening instrument must be demonstrated through empirical evidence of the advantages, usefulness, or validity of such an instrument. Legislation is constantly being influenced by court decisions involving unfair practice and discrimination suits (Institute of Industrial Relations, 1971). Due to the increased national concern for more effective police forces, and the increased interest by the federal government in providing equitable opportunities for minority candidates to function effectively in law enforcement agencies, it is imperative that the performance of both majority and minority candidates and officers be assessed.

The difficulty in formulating scientific performance appraisal systems for the law enforcement profession stems from

the lack of a congruent set of standards and criteria with which to assess performance. Many of the recurrent difficulties inherent in police selection and performance studies arise from the differences of opinion regarding the definition of a responsive, effective police officer (Kelly and Farber, 1974), and from the variance in the criteria employed in judging "good" police performance (Brodsky, 1973). Limits are imposed on the standardization of performance measures. This problem stems from the fact that the police officer is continually confronted with situations involving unique or unusual circumstances and contingencies in daily work schedules. Additionally, work is performed somewhat independently because the officer generally has discretion concerning the manner in which he/she distributes his/her time across various duties (Robinson, 1970a).

Job functions of the police officer tend to be vague because police work generally does not involve output dimensions in terms of standard measurable quantities. While police agency effectiveness is commonly assessed in terms of a variety of productivity measures such as crime rate statistics, victimization surveys, stolen goods recovered, and personnel complaints registered, it is much more difficult to determine the effectiveness of the individual patrol officer. A sole reliance on productivity measures may predispose an officer to overreact and overemphasize such performance objectives to the extent of disregarding aspects of police work not associated with productivity (National Advisory Commission on Criminal Justice

Standards and Goals, 1973; Mills, 1972). Such a production criterion as number of arrests is inappropriate as a measure of success in crime control unless such factors as the qualitative aspects of the manner of arrest, the contingent circumstances of the arrest, and the ultimate disposition of the case are considered. Quantitative production standards may also distort performance evaluation by ignoring such variables as officer discretion, courtesy, impartiality, and responsiveness to the community in terms of general protection, counseling, referral, and service (President's Commission on Law Enforcement and Administration of Justice, 1967).

Fluctuations in the conventional productivity indicators (e.g., number of arrests made, number of traffic citations issued, number of radio calls answered) used in police evaluation can sometimes be justified by the task variations and situational differences surrounding the work assignment of each officer. The opportunities to satisfy production standards vary in accordance with such factors as geographical patrol area, work shift, and priority duties assigned. Thus, incomplete and sometimes distorted conceptions of the effectiveness of crime control, community protection, and services provided by the police may result from reliance on single or incompatible combinations of performance measures.

The basic dilemma, therefore, in securing adequate, observable criteria for the police profession arises from the intricate nature of police work. Part of the problem in

establishing criteria revolves around the inability or indifference of police administrators in determining the methods and techniques which should be used in patrol activities (Wilson, 1968). Many police administrators take the attitude that the best which might be hoped for is a rough evaluation of compliance to the police role, or an overall rating of the police officer on "quality of work" or "overall suitability" (Landy and Farr, 1975). Adequate performance appraisal is also limited by the indirect supervision under which the officer works. Since the delivery of non-enforcement services constitutes a significant part of the policing function, more sophisticated supervision techniques and appropriate quality control measures should be developed to ensure that the effectiveness of the individual officer as well as that of the police agency can be ascertained in responding to total community needs and expectations (National Advisory Commission on Criminal Justice Standards and Goals, 1973; Germann, 1971).

In conclusion, while there is an increasing concern by government and several community sectors in improving the quality of police personnel, there is a noticeable scarcity of comprehensive and carefully designed studies to determine job components, to establish performance criteria, and to determine predictive indices of police success and failure (Landy and Farr, 1975; Hogan, 1972; Baehr, Furcon, and Froemel, 1968). Personnel qualifications and performance

standards cannot be established through the indiscrete use of personality factors based upon certain presumptions and intuitions. Rather, such requirements and standards must be established through a more objective investigation of task elements and their relationship to individual characteristics. Although there is an urgent need for identifying reliable and valid predictors of police performance (Azen, Snibbe, and Montgomery, 1973), the vast majority of police studies have been non-validity studies. Numerous governmental reports and psychological studies emphasize the need for systematic research aimed at determining the relationships among predictors and criteria in the field of law enforcement if the police system is to improve. In essence, the research literature suggests that neither the predictors nor criteria have been clearly identified and related in any complex prediction model utilizing relative contributions of multiple variables in police research. Intensive research is needed in the specific areas of assessing the competence of police officers in various components of police work; determining the validity and reliability of selection, training, and performance measures; and determining relationships of predictor-criterion measures for subgroups differentiated according to such variables as race, sex, and mental ability.

Police effectiveness as a general concept must be evaluated in terms of proficiency in the constituent aspects of police work. A job analysis is an essential part of an

evaluative program if a determination of the work tasks and behavioral dimensions of the patrol officer's job is to be accomplished (Atlanta Regional Commission, 1974; Baehr, Furcon, and Froemel, 1968). An effective training program must be based on a complete and accurate knowledge of the tasks of the patrol officer (Allen, 1976; Badalamente, et al., 1973; McManus, et al., 1970; Robinson, 1970b). The City of Atlanta Personnel Office, in conjunction with the Planning and Research Section of the Atlanta Police Department, conducted a task analysis of the department's Patrol Section in 1974. An explanation of the manner in which the task analysis was conducted and the results obtained are provided in Appendix A.

Study Objectives

The present investigation focuses on several salient variables in the personnel research literature. It is an attempt to validate a mental ability selection instrument against training and job performance; to validate training performance against job performance; and to perform comparative tests for the demographic variables of race, sex, and mental ability on the performance measures. The research was conducted through cooperative arrangements with the Atlanta Police Department and the City Personnel Office.

The specific objectives of this study were: (a) to determine whether significant race, sex, and/or mental ability

differences exist on the police academy training variables for the sample; (b) to determine whether significant race, sex, and/or mental ability differences exist on the supervisory rating variables for the sample; (c) to determine whether significant race, sex, and/or mental ability differences exist on the supervisory rating variables for the sample when the variance on police academy training variables is controlled by covariance adjustments; (d) to determine whether the police academy training variables intercorrelate significantly for the total pooled sample; (e) to determine whether the supervisory rating variables intercorrelate significantly for the total pooled sample; (f) to determine which police academy training variables, if any, are significantly related to which supervisory rating variables for the total pooled sample; (g) to determine which police academy training variables and supervisory rating variables, if any, are significantly related to the mental ability variable for the total pooled sample; and (h) to determine whether the set of police academy training variables is significantly related to the set of supervisory rating variables for the total pooled sample.

Specific Hypotheses

Hypothesis 1: No significant race differences will occur on the police academy training variables on which the subjects are compared.

Hypothesis 2: No significant sex differences will occur on the police academy training variables on which the subjects are compared.

Hypothesis 3: No significant mental ability differences will occur on the police academy training variables on which the subjects are compared.

Hypothesis 4: No significant race differences will occur on the supervisory rating variables on which the subjects are compared.

Hypothesis 5: No significant sex differences will occur on the supervisory rating variables on which the subjects are compared.

Hypothesis 6: No significant mental ability differences will occur on the supervisory rating variables on which the subjects are compared.

Hypothesis 7: No significant race differences will occur on the supervisory rating variables when police academy training variables are treated as covariates.

Hypothesis 8: No significant sex differences will occur on the supervisory rating variables when police academy training variables are treated as covariates.

Hypothesis 9: No significant mental ability differences will occur on the supervisory rating variables when police academy training variables are treated as covariates.

Hypothesis 10: No overall significant interactions will occur on any of the multivariate factorial analyses using the factors

of race, sex, and mental ability.

Hypothesis 11: The set of police academy training variables will be significantly related to the set of supervisory rating variables for the total sample.

CHAPTER II

METHODOLOGY

Subjects

The sample consisted of 227 police officers from the Patrol Section of the Field Operations Division, the Atlanta Police Department. All subjects had the rank of patrol officer. Information concerning name, race, sex, employment data, zone assigned, and work shift was procured from the Personnel Section of the Atlanta Police Department. With regard to race, the personnel records classified each subject as either black or white. The subgroup sizes, based on race and sex, consisted of 147 white males, 51 black males, 17 white females, and 12 black females.

Variables and Instruments

A Mental Ability Instrument

Scores on the Advanced Level of the Otis-Lennon Mental Ability Test served as a measure of mental ability. The Otis-Lennon Mental Ability Test had been chosen by the Atlanta Police Department and City Personnel Office as an initial screening instrument for police officers. The test had been administered to police applicants by the City of Atlanta Personnel Office, and scores for the sample were obtained from the personnel records. Description, explanation, and

psychometric properties of the test may be found in Otis and Lennon (1967) and Buros (1972). The criterion score selected as a passing mark for a police applicant was 35, with the maximum possible score being 80. If an applicant failed the test, he/she retook the alternate form of the test after a minimal period of six months. There was no limit to the number of times an applicant could take the test.

The sample was differentiated according to measured mental ability. The availability of Otis-Lennon test scores permitted assignment of individuals to either the lower (i.e., score range of 35 to 48) or the higher mental ability group (i.e., score range of 49 to 80) based upon a sample-determined median score. The resulting subgroup totals were 114 and 113 respectively.

Police Academy Training School Variables

Training scores for the subjects were obtained from the Atlanta Police Academy and the City of Atlanta Personnel Office. Sixteen successive police recruit training classes were samples. Although there have been slight modifications in the topical areas and training school variables over the thirty-seven months which these classes have spanned, only the thirteen variables which were common to each of the sixteen classes were used. The thirteen training school variables used in this study were: (A) Accident Investigation; (B) Traffic Ordinances; (C) State Traffic Laws; (D) Speed and Skidmarks; (E) City Ordinances; (F) Evidence; (G) State

Laws; (H) Report Writing; (I) Rules and Regulations; (J) First Aid; (K) Notebook; (L) Practical Pistol Course; and (M) Classroom Average. A description of these training school variables is provided in Appendix B.

Supervisory Rating Variables

The instrument used to obtain measures of performance for the patrol officers in the sample was a set of supervisory rating scales developed by Landy and Farr (1973, 1975). The ratings were based on a nine-point behaviorally-anchored scale which also had anchors midway between the numerical points of the scale, and therefore constituted a seventeen-point scale.

A scale was provided for each of the following eight trait and performance dimensions: (1) "Job Knowledge," defined as awareness of procedures, laws, and court rulings, and changes in them; (2) "Judgment," defined as observation and assessment of the situation, and taking appropriate action; (3) "Initiative," defined as individual personal performance conducted without either direct supervision or commands, including suggestions for improved departmental procedures; (4) "Dependability," defined as predictable job behavior, including attendance, promptness, and reaction to boredom, stress, and criticism; (5) "Demeanor," defined as professional bearing as determined by overall neatness of uniform, personal grooming, and general physical condition; (6) "Attitude," defined as general orientation toward the law enforcement

profession and the department; (7) "Relations with Others," defined as the ability to deal with people he/she comes into contact with during the performance of his/her job, including the public, fellow officers, and supervisory personnel; and (8) "Communication," defined as the ability to make oneself understood and gather and transmit information, both in oral and written fashion.

An explanation of the manner in which the scales were constructed, and the psychometric properties of the scales may be found in Appendix C. The performance description scales composing the supervisory rating instrument are contained in Appendix D. The instruction and example booklet for the performance description scales is depicted in Appendix E. The supervisory officers who rated the patrol officers were instructed to read this booklet.

The reliability of the supervisory rating instrument could not be directly determined since it was not feasible to obtain more than one supervisory rating form for each patrol officer in the sample. However, Landy and Farr (1975) investigated the psychometric properties of the rating scales and found that the instrument does have substantial reliability. They note that "the reliabilities are sufficiently high and the intercorrelations sufficiently low to justify the conclusion that each scale taps some unique portion of systematic rating variance." The construct validity of the instrument was also investigated by determining

relationships between the instrument and a variety of performance measures, both subjective and objective, obtained in numerous police agencies included in the national survey. Landy and Farr concluded from this investigation that "the general results indicated that the (supervisory) scales appeared to be measuring various components of the job performance of police patrol officers."

Police supervisors of the Atlanta Police Department did not participate in the construction of the supervisory rating instrument. Although it is doubtful that participation would have altered the final format of the rating instrument, it may have contributed to more cooperation and conscientiousness on the part of the supervisors in rating their subordinates. However, Landy, et al. (1976) have reported that, on the basis of data gathered in 58 municipal police agencies, many agencies were able to use the supervisory scales effectively despite the fact that raters from those agencies were not directly involved in the development of the rating scales.

Although several researchers (e.g., Anastasi, 1976; Whisenand, 1976; Borman and Dunnette, 1975; Iannone, 1975; Guion, 1965; Beck, 1961; Guilford, 1954) have advocated rater training as a means for improving evaluative accuracy of ratings, it was not feasible to provide training for the supervisory police officers who rated the patrol officers in this study. Studies (e.g., Borman, 1975; Latham, Wexley,

and Pursell, 1975; Driver (unpublished) in McCormick and Tiffin, 1974) have indicated that rater training reduces the likelihood of substantial rater error, and thus provides more accuracy in performance assessment. However, the use of behaviorally-anchored rating scales tends to reduce potential rater error by infusing more objectivity into the rating process (Landy and Trumbo, 1976; Campbell et al., 1973; Dunnette, 1966; Peters and McCormick, 1966; Smith and Kendall, 1963; Barrett, et al., 1958).

Procedure

Data were collected through the precinct headquarters of the five patrol zones of the City of Atlanta. These zones represent approximately equal geographic territories of the metropolitan Atlanta area.

The experimenter visited each of the zone captains with a list, categorized according to watch shift, of names of patrol officers included in the sample and assigned to the particular zone. For each zone precinct, a list had been compiled from a master file of patrol officers obtained from personnel cards on file in the Personnel Section of the Atlanta Police Department. An appropriate number of supervisory rating forms and instruction and example booklets were distributed to the zone precincts.

The captain was instructed to have the supervisor (i.e., lieutenant or sergeant) who was most familiar with the job performance of each patrol officer on the list to read the

instruction and example manual, and rate the officer(s) on the eight dimensions of performance using the supervisory rating form.

A total of 28 supervisors participated in rating the patrol officers. The demographic variables of rank, race, and sex were obtained for the supervisory personnel who contributed at least one rating form to the study. One supervisory rating was obtained for each of the subjects in the sample.

Experimental Design and Data Analyses

A possible source of individual police officer assessment contamination which has been somewhat neglected in police performance research is the job experience variable. In an attempt to control for wide variations in job experience, the sample was limited to subjects initially employed as police officers within a three-year span (i.e., from January 1972 through December 1974).

The experimental design of the study provided for a further control of the experience factor. The length of the training period was an extraneous variable which has been statistically controlled (i.e., adjusted) through covariance analysis. Recruit Training Classes #65 through #75 were six-week training programs, while Recruit Training Classes #76 through #80 were thirteen-week training programs. A total of 121 subjects participated in the six-week program, and a total of 106 subjects participated in the thirteen-week

program. Six weeks of the police academy training curriculum is a mandatory requirement of the State of Georgia, and represents the minimal amount of training necessary for a municipality to give legal sanction to police officers in law enforcement activities. The six-week training curriculum was abolished in December 1973. Seven weeks were added to the state's requirements in January 1974. The use of "length of training period" as a covariate served as an indirect control for job experience because the subjects who participated in the six-week training program had more job experience since they were employed earlier than the subjects who participated in the thirteen-week training program.

Five a priori and two a posteriori analyses were performed on the data. A schematic representation of the experimental design for the study may be found in Table 2. The statistical techniques used in the analyses were univariate analysis of variance (Analysis Ia), multivariate analysis of variance (Analyses I, II, IIa, and III), and correlational analysis (Analysis IV and V).

The first three a priori analyses utilized a fixed-effects, nonorthogonal 2×2×2 multivariate factorial design such that two levels were assigned to each of three factors. The independent variables of race, sex, and mental ability served as factors in this design. The factor "race" had levels black and white; the factor "sex" had levels male and female; and the factor "mental ability" had levels higher

Table 2. Schematic Representation of the Experimental Design

ANALYSIS	HYPOTHESIS	VARIABLES		
		Independent	Dependent	Covariate
I	H1, H2, H3, H10	Race, Sex, Mental Ability	Training School Scores	Length of Training Period
I(a)		Race, Sex	Mental Ability Scores	
II	H4, H5, H6, H10	Race, Sex, Mental Ability	Supervisory Ratings	Length of Training Period
II(a)		Race of Subordinate, Race of Supervisor	Supervisory Ratings	
III	H7, H8, H9, H10	Race, Sex, Mental Ability	Supervisory Ratings	Length of Training Period, Training School Scores
IV			Mental Ability Scores Training School Scores, Supervisory Ratings	
V	H11		Training School Scores, Supervisory Ratings	

mental ability group and lower mental ability group. The first a priori multivariate factorial design (Analysis I) treated the thirteen training school subject areas as dependent variables, and treated length of training period as a covariate. The second a priori multivariate factorial design (Analysis II) treated the eight supervisory rating scales as dependent variables, and treated length of training period as a covariate. The third a priori multivariate factorial design (Analysis III) treated the eight supervisory rating scales as dependent variables, and treated length of training period and the thirteen training school subject areas as covariates. Analysis IV involved obtaining an intercorrelation matrix for the total pooled sample of patrol officers. Twenty-two measures (i.e., the thirteen training school subject areas, the eight supervisory rating scales, and the Otis-Lennon Mental Ability Test) served as dependent variables. Analysis V involved obtaining canonical correlations to test the relationship between the set of training school variables and the set of supervisory rating variables for the total pooled sample of patrol officers.

The first a posteriori analysis (Analysis Ia) utilized a fixed-effects, nonorthogonal 2×2 univariate factorial design with the factors of race and sex of the patrol officers being used as independent variables, and mental ability scores serving as a dependent variable. The second a posteriori analysis (Analysis IIa) utilized a fixed-effects, non-

orthogonal 2×2 multivariate factorial design with the factors of race of patrol officers and race of the supervisors who rated the patrol officers being used as independent variables, and the eight supervisory rating scales serving as dependent variables.

The conceptualization for multivariate analysis of variance focuses upon a vector containing several dependent variables with observations on each of these variables for each of the subjects in the design. While analysis of variance operates with means on individual variables, multivariate analysis of variance analyzes vectors of means where each element of the vector is a group's mean for a particular variable. Each group's vector of mean scores is termed that group's "centroid."

The purpose of the multivariate analysis of variance technique is to determine whether statistically significant differences exist between two or more groups based upon the groups' members' scores on the set of dependent variables (Amick and Crittenden, 1975). Multivariate analysis of variance, as applied to the factorial experimental design, is a technique which derives linear combinations of the dependent variables which best differentiate among the treatment levels of the independent variables (i.e., factors). Each of these linear combinations is tested for significance with multivariate tests (e.g., Wilks' lambda criterion) for each factor and interaction effect of the design. Multivariate analysis

of variance is appropriate for Analyses I, II, IIa, and III because the objectives for these analyses included determination of whether significant race, sex, and/or mental ability differences exist on sets of dependent variables.

Covariance adjustment of length of training period (Analysis I, II, and III) and the thirteen training school variables (Analysis III) allows for control of the effects of training performance. A covariate is a variable which is related to (i.e., covaries with) the predictor and/or criterion variables. Covariance adjustments allow for the removal of potential sources of bias from the experiment through statistical control rather than through experimental control (Harris, 1975). The procedure involves measuring one or more concomitant variates in addition to the dependent variates. The concomitant variate represents a source of variation which might affect the dependent variates. Length of training period is treated as an extraneous variable in the initial three a priori analyses. The training school subject areas are treated as extraneous variables in Analysis III. Through analysis of covariance, the dependent variates can be adjusted to remove the effects of the uncontrolled source of variation represented by the concomitant variate (Kirk, 1968).

Multivariate analysis of covariance is an extension of univariate analysis of covariance. It consists of a multivariate analysis of variance performed on a set of corrected

scores, each corrected score having resulted from a separate regression analysis of the relationship between the covariates and that particular outcome measure (Harris, 1975).

The three a priori multivariate factorial analyses were performed to test the first ten hypotheses of this study. Factors in the analyses are race (referred to as R), sex (referred to as S), and mental ability (referred to as M). The model equation tested in these three analyses is:

$$\underline{Y}_{ijk} = \underline{\mu} + \underline{R}_i + \underline{S}_j + \underline{M}_k + \underline{RS}_{ij} + \underline{RM}_{ik} + \underline{SM}_{jk} + \underline{RSM}_{ijk} + \underline{e}$$

where $\underline{Y}_{ijk}$ is an n×1 random vector of scores for an observation on $\underline{n}$ variables in the i,j,k cell of the design, $\underline{\mu}$ is an n×1 vector of constants, $\underline{R}_i$ is an n×1 vector of constants for the n dependent variables on the $\underline{i^{th}}$ level of factor R, $\underline{S}_j$ is an n×1 vector of constants for the $\underline{n}$ dependent variables on the $\underline{j^{th}}$ level of factor S, $\underline{M}_k$ is an n×1 vector of constants for the $\underline{n}$ dependent variables on the k^{th} level of factor M, $\underline{RS}_{ij}$ is an nx1 vector of constants for the i,j level combination of $\underline{n}$ dependent variables, $\underline{RM}_{ik}$ is an n×1 vector of constants for the i,k level combination of $\underline{n}$ dependent variables, $\underline{SM}_{jk}$ is an n×1 vector of constants for the j,k level combination of $\underline{n}$ dependent variables, $\underline{RSM}_{ijk}$ is an n×1 vector of constants for the i,j,k level combination of $\underline{n}$ dependent variables, and $\underline{e}$ is an n×1 vector of $\underline{n}$ random error variables. Each of the terms in the model equation indicates a vector

of parameters (see Timm, 1975).

The purpose of the a priori multivariate factorial analyses was to determine whether the complete model equation or some simplification of it represented the true state of affairs for the sample. A mathematical model is associated with every experimental design, and this model purports to include all sources of variability affecting individual scores on the dependent variable(s). To the extent that the model accurately represents these sources of variability, the effects of a treatment or treatments can be evaluated (Kirk, 1968).

Cramer's (1973) MANOVA program was used to analyze the multivariate factorial designs (Analysis I, II, IIa, and III) and the univariate factorial design (Analysis Ia). This program performs univariate and multivariate analysis of variance, covariance, and regression, and provides an exact solution in either the orthogonal or nonorthogonal case. For the multivariate analysis of variance, the overall test of significance is performed using Wilks' lambda criterion, often referred to as the "ratio of determinants test" (Amick and Crittenden, 1975). Wilks' lambda computes the ratio of the "likelihood" of the data under the assumption of population mean vectors identical to the corresponding sample mean vectors, to the "likelihood" of the observed data under the assumption of no differences among population mean vectors (Harris, 1975). That is, it is a statistic for testing the significance of the overall difference among several sample centroids,

constituting a multivariate extension of the F-ratio test in simple analysis of variance (Tatsuoka, 1971). The test of group differences allows the determination of whether the group centroids are significantly different (Amick and Crittenden, 1975). While Cramer's multivariate analysis of variance program uses Wilks' lambda criterion to perform the overall test of significance, it also uses Rao's (1952) statistic, which is a function of lambda that approximately follows an F distribution, to report probability levels (Cramer, 1973).

In addition to an overall test of significance for the model effects, the MANOVA program performs univariate F tests for each dependent variable. When the test of lambda produces an F value that allows the researcher to reject the null hypothesis of equal group centroids, the univariate F ratios for the analysis of variance may be examined to determine which elements of the vector variable are critical to the difference(s) between or among the groups (Bock and Haggard, 1968). Therefore, a univariate, or single degree of freedom, F test is performed for each of the dependent variables involved in the overall multivariate test of significance to provide some indication of the degree to which each of the variables contributes to the overall significant effect. The results of the univariate tests may facilitate the interpretation of the overall significance test result.

An alpha level of .05 was chosen as the criterion of

statistical significance for the Wilks' lambda test of significance. A nominal alpha level of .05 was also chosen as the criterion of statistical significance for the univariate F tests. However, the univariate tests in the multivariate design represent a family or composite of tests which, if each were tested at the .05 alpha level, would have an inflated alpha level actually higher than the nominal alpha level of .05. Therefore, in order to reduce the probability of Type I error (i.e., rejecting the null hypothesis when it is true), the nominal alpha level was divided by the number of univariate tests to be performed. Thus, for the multivariate analysis using thirteen dependent variables (Analysis I), the critical alpha level for the univariate tests was calculated to be .004. For the multivariate analyses using eight dependent variables (Analysis II, IIa, and III), the critical alpha level for the univariate tests was calculated to be .006. While the nominal alpha level (.05) represents the significance level for the family of univariate tests, the critical alpha level represents the significance level for each univariate F test in the multivariate analyses.

A nonorthogonal design was used in the factorial analyses. The term "nonorthogonal" refers to an experimental design in which the numbers of observations are not equal in each cell of the design. This indicates that the estimated parameters within a model are statistically dependent of one another (Appelbaum and Cramer, 1974). The problem with the

nonorthogonal design is essentially one of adjusting for the effects of the confounding factor(s) (Appelbaum, 1974).

The significance test results reported in this study represent tests of the model effects which have eliminated (i.e., adjusted for) the confounding effects of other factors when they are present. For example, in the tests of each of the main effects (i.e., race, sex, and mental ability) in Analyses I, II, and III, the other two factors have been eliminated. The reported results for the race effect, for instance, are not influenced or biased by the effects of sex and mental ability. For the tests of each of the first-order interaction effects (i.e., race by sex (R×S), race by mental ability (R×M), and sex by mental ability (S×M)), the other two interaction effects and the main effects have been eliminated. The reported results for the R×S interaction effect, for instance, are not influenced or biased by the interaction effects R×M and S×M, or the main effects of race, sex, and mental ability. For the test of the second-order interaction effect (i.e., race by sex by mental ability (R×S×M)), all first-order interaction effects and main effects have been eliminated.

Thus, nonorthogonal designs necessitate special analyses because of the unequal cell frequency problem. The method of adjusting for effects in the design is necessary to estimate treatment effects without bias and to provide unbiased tests of significance (Appelbaum and Cramer, 1974).

Analysis IV involved obtaining an intercorrelation matrix, based on Pearson product-moment correlations, for the total pooled sample of patrol officers. Twenty-two measures (i.e., the Otis-Lennon Mental Ability Test, the training school subject areas, and the supervisory rating scales) were used as dependent variables. The intercorrelational analysis portrays the inter-individual relationships among the various performance measures. This analysis used Fisher's r to Z transformation, critical values of an approximation to Student's t statistic, and the multistage Bonferroni procedure to determine the significant correlation coefficients in the matrix. An alpha level of .05 was chosen as the criterion of statistical significance for the correlation coefficients.

Significance of the correlation coefficients was tested using a modification of the Bonferroni procedure suggested by Larzelere and Mulaik (1977). They propose a technique which they call a "multistage Bonferroni procedure" to correct for the unnecessary stringency of the conventional Bonferroni procedure. With this procedure, the probability of making at least one Type I error in a family of tests is controlled at a level less than or equal to the nominal familywise probability. The probability of making one or more Type I errors on any subset of tests is no greater than the familywise probability. In essence, the multistage Bonferroni procedure controls the probability of making at least one Type I

error in a family of tests or in any subset of that family, while providing more power than the traditional, one-stage Bonferroni procedure.

Morrison (1976) discusses a Bonferroni procedure for multiple tests of correlation using Fisher's r to Z transformation. Fisher's r to Z transformation is necessary for obtaining confidence intervals and for testing against any nonzero null hypothesis. Use of the Bonferroni t-test procedure often requires computation of "odd" critical values of Student's t distribution (Harris, 1976). In Analysis IV, the critical value for this test was approximated through a formula, provided by Zelen and Severo (1964), which calculates a two-tailed critical value, for the alpha level, of the unit normal distribution based upon a logarithmic function of the t value. This value and the standard error of Z were used in the calculation of the critical value for Fisher's r to Z transformation using a logarithmic function. The critical correlation coefficient was calculated based on a conversion of the formula used in obtaining the critical value for Fisher's r to Z transformation. The critical value for Fisher's r to Z transformation represents the critical unit normal deviate approximation for the per-test alpha level. The critical value for Fisher's r to Z transformation entered into the final calculation.

Analysis V involved testing the relationship between the set of training school variables and the set of

supervisory rating variables for the total pooled sample of patrol officers. This relationship was tested using canonical correlation analysis. Canonical correlation is the extension of linear multiple regression to the case of multiple criterion variables. It is appropriate for the situation in which a set of continuous predictor variables is to be related to a set of continuous criterion variables (Weiss, 1976). For this study, the set of continuous predictor variables consisted of the training school subject areas, and the set of continuous criterion variables consisted of the supervisory rating scales. The basic purpose of canonical correlation analysis is to derive a linear combination from each of the sets of variables in such a way that the correlation between the two linear combinations is maximized (Nie, et al., 1975). The canonical correlation is the product-moment correlation between the variate of the predictor variables and the variate of the criterion variables when both variates are derived from beta weights chosen to maximize that correlation. The criterion variables and the predictor variables are weighted simultaneously, by means of two sets of regression weights, to arrive at two variates which correlate as highly as possible with each other (Weiss, 1976). Therefore, for canonical correlation, the problem is reduced to finding linear combinations of the predictors and criteria that will maximize the relationship between the two sets of variables to provide an overall measure of their relationship

(Harris, 1975; Anderson, 1966). Usually the researcher is interested in the maximal predictive relationship between two sets of variables, in which case, the first or maximum canonical correlation is appropriate (Weiss, 1976). An alpha level of .05 was chosen as the criterion of statistical significance for the canonical correlations.

All analyses were performed using the CYBER 70 Model 74-28/6400 Multi-Processor Computer System. The multivariate factorial analyses and univariate factorial analysis were performed by inputing control and data cards into an express job reader. For Analysis IV, correlation coefficients were obtained from a program provided by Nie, et al. (1975) in the Statistical Package for the Social Sciences. Critical correlation values for testing significance of the coefficients were calculated using a fortran program punched into a computer terminal. Analysis V was performed using the canonical correlation program also provided by Nie, et al. (1975). Control and data cards were used as input for the correlational analyses.

CHAPTER III

RESULTS

Preliminary Manipulations and Analyses

The descriptive statistics which were obtained for the data give an indication of the overall level of subgroup and total sample performance (group means) and the variability of scores for each variable (standard deviation). Means and standard deviations were calculated for the four subgroups (black females, Table 3; white females, Table 4; black males, Table 5; and white males, Table 6) of patrol officers on twenty-two variables (i.e., Otis-Lennon Mental Ability Test scores, thirteen training school subject areas, and eight supervisory rating scales). Also, means and standard deviations were calculated for the four subgroups differentiated according to mental ability level (i.e., higher mental ability or lower mental ability) within each subgroup. These normative statistics were obtained for twenty-one variables (i.e. training school subject areas and supervisory rating scales). Finally, descriptive statistics were obtained on the twenty-two variables for the total pooled sample of patrol officers (Table 7).

Table 3. Means (M) and Standard Deviations (SD) of the Variables for Black Female Patrol Officers

VARIABLE	TOTAL SAMPLE*		HIGHER MENTAL ABILITY**		LOWER MENTAL ABILITY***	
	M	SD	M	SD	M	SD
Otis-Lennon Mental Ability	44.33	7.64				
Accident Investigation	85.33	8.26	81.00	13.53	86.78	6.22
Traffic Ordinances	80.33	8.08	81.67	5.69	79.89	8.99
State Traffic Laws	87.58	10.90	82.00	20.88	89.44	6.23
Speed and Skidmarks	76.08	6.32	82.67	4.16	73.89	5.37
City Ordinances	67.75	8.35	71.00	8.54	66.67	8.50
Evidence	79.17	10.18	83.00	7.21	77.89	11.05
State Laws	90.50	11.40	93.67	7.77	89.44	12.59
Report Writing	82.33	6.30	84.33	4.04	81.67	6.96
Rules and Regulations	82.83	7.81	83.67	2.08	82.56	9.08
First Aid	83.17	8.17	89.67	5.77	81.00	7.89
Notebook	94.67	5.76	93.00	7.81	95.22	5.38
Practical Pistol Course	71.92	6.79	70.67	9.24	72.33	6.42
Classroom Average	82.08	4.17	84.67	4.04	81.22	4.06
Job Knowledge	5.79	1.78	6.33	2.08	5.61	1.76
Judgment	5.75	2.21	6.83	2.47	5.39	2.15
Initiative	5.63	2.33	4.33	3.51	6.06	1.88
Dependability	5.58	2.24	5.17	2.25	5.72	2.36
Demeanor	6.33	1.91	6.83	1.61	6.17	2.06
Attitude	5.88	1.98	5.67	2.08	5.94	2.07
Relations with Others	6.58	1.79	6.83	2.02	6.50	1.84
Communication	6.21	1.62	6.50	1.32	6.11	1.76

*$\underline{N}$ of 12
**$\underline{N}$ of 3
***$\underline{N}$ of 9

Table 4. Means (M) and Standard Deviations (SD) of the Variables for White Female Patrol Officers

VARIABLE	TOTAL SAMPLE*		HIGHER MENTAL ABILITY**		LOWER MENTAL ABILITY***	
	M	SD	M	SD	M	SD
Otis-Lennon Mental Ability	56.06	10.91				
Accident Investigation	91.06	7.90	93.92	3.70	84.20	11.35
Traffic Ordinances	88.76	6.09	89.00	5.24	88.20	8.50
State Traffic Laws	88.71	7.10	88.33	7.46	89.60	6.84
Speed and Skidmarks	86.94	10.15	87.83	11.19	84.80	7.69
City Ordinances	84.18	9.85	86.50	8.34	78.60	11.91
Evidence	82.00	10.40	81.33	11.99	83.60	5.77
State Laws	89.12	7.78	92.17	5.69	81.80	7.63
Report Writing	88.41	4.61	88.58	5.09	88.00	3.67
Rules and Regulations	85.71	4.12	87.75	2.67	80.80	2.39
First Aid	87.82	4.42	87.92	4.34	87.60	5.13
Notebook	98.24	4.02	97.83	4.76	99.20	.84
Practical Pistol Course	71.53	9.38	73.83	9.50	66.00	7.04
Classroom Average	87.65	3.06	88.83	2.41	84.80	2.68
Job Knowledge	5.97	1.48	5.83	1.68	6.30	.91
Judgment	5.91	1.99	5.79	2.25	6.20	1.30
Initiative	6.15	2.21	5.96	2.48	6.60	1.52
Dependability	5.71	2.18	5.75	2.46	5.60	1.52
Demeanor	6.09	1.82	5.92	1.93	6.50	1.66
Attitude	5.97	1.62	5.88	1.86	6.20	.91
Relations with Others	6.50	1.76	6.50	1.94	6.50	1.41
Communication	6.82	1.55	7.25	1.54	5.80	1.10

*$\underline{N}$ of 17
**$\underline{N}$ of 12
***$\underline{N}$ of 5

Table 5. Means (M) and Standard Deviations (SD) of the Variables for Black Male Patrol Officers

VARIABLE	TOTAL SAMPLE*		HIGHER MENTAL ABILITY**		LOWER MENTAL ABILITY***	
	M	SD	M	SD	M	SD
Otis-Lennon Mental Ability	42.78	6.80				
Accident Investigation	80.98	10.21	85.36	8.25	79.78	10.45
Traffic Ordinances	77.31	7.67	80.09	7.08	76.55	7.74
State Traffic Laws	81.02	10.05	81.82	12.46	80.80	9.46
Speed and Skidmarks	78.29	10.74	81.27	10.58	77.48	10.78
City Ordinances	70.59	9.61	73.64	11.12	69.75	9.13
Evidence	78.51	13.04	79.73	15.15	78.18	12.60
State Laws	79.04	10.90	83.64	11.11	77.78	10.63
Report Writing	75.22	10.57	77.36	3.80	74.63	11.74
Rules and Regulations	80.84	6.98	82.91	6.91	80.28	6.98
First Aid	80.78	6.74	84.91	6.55	79.65	6.42
Notebook	88.73	6.99	90.46	9.20	88.25	6.32
Practical Pistol Course	74.29	9.85	75.46	8.70	73.98	10.22
Classroom Average	78.24	4.50	80.46	5.18	77.63	4.16
Job Knowledge	6.46	1.75	6.86	1.32	6.35	1.85
Judgment	6.42	1.89	6.73	1.95	6.34	1.89
Initiative	6.28	1.92	6.82	1.83	6.14	1.94
Dependability	6.23	2.30	6.68	1.81	6.10	2.42
Demeanor	6.53	1.98	7.23	1.94	6.34	1.97
Attitude	6.64	1.95	7.09	1.28	6.51	2.09
Relations with Others	6.75	1.92	7.41	1.34	6.56	2.03
Communication	6.24	2.27	7.27	1.40	5.95	2.39

*$\underline{N}$ of 51
**$\underline{N}$ of 11
***$\underline{N}$ of 40

Table 6. Means (M) and Standard Deviations (SD) of the Variables for White Male Patrol Officers

VARIABLE	TOTAL SAMPLE*		HIGHER MENTAL ABILITY**		LOWER MENTAL ABILITY***	
	M	SD	M	SD	M	SD
Otis-Lennon Mental Ability	52.31	11.71				
Accident Investigation	84.24	9.43	84.71	9.33	83.55	9.61
Traffic Ordinances	80.99	8.48	81.95	8.04	79.60	8.98
State Traffic Laws	83.98	10.10	85.31	9.83	82.05	10.27
Speed and Skidmarks	85.28	9.38	87.16	9.17	82.55	9.08
City Ordinances	78.07	10.65	79.55	8.68	75.92	12.77
Evidence	81.95	11.01	84.12	9.90	78.80	11.84
State Laws	81.77	9.08	83.33	7.98	79.50	10.11
Report Writing	80.69	10.75	82.15	10.69	78.58	10.56
Rules and Regulations	85.67	7.00	87.21	6.86	83.45	6.63
First Aid	85.52	5.67	86.75	5.19	83.73	5.89
Notebook	91.50	9.27	91.23	9.87	91.90	8.39
Practical Pistol Course	78.73	9.85	79.93	9.95	77.00	9.53
Classroom Average	82.69	5.04	84.16	4.86	80.57	4.54
Job Knowledge	6.79	1.62	6.89	1.68	6.64	1.53
Judgment	6.52	1.79	6.59	1.83	6.43	1.75
Initiative	6.60	1.94	6.54	2.00	6.68	1.85
Dependability	6.75	2.04	6.67	2.18	6.88	1.82
Demeanor	6.67	1.91	6.76	1.93	6.53	1.90
Attitude	6.57	1.88	6.51	2.03	6.67	1.66
Relations with Others	6.92	1.60	6.93	1.68	6.91	1.48
Communication	7.02	1.73	7.16	1.73	6.83	1.72

*$\underline{N}$ of 147
**$\underline{N}$ of 87
***$\underline{N}$ of 60

Table 7. Means and Standard Deviations of the Variables for the Total Sample of Patrol Officers

VARIABLE	MEAN	STANDARD DEVIATION
Otis-Lennon Mental Ability	50.03	11.39
Accident Investigation	84.07	9.71
Traffic Ordinances	80.71	8.54
State Traffic Laws	83.86	10.10
Speed and Skidmarks	83.35	10.18
City Ordinances	76.30	11.04
Evidence	81.03	11.44
State Laws	82.17	9.99
Report Writing	80.13	10.67
Rules and Regulations	84.44	7.13
First Aid	84.50	6.34
Notebook	91.55	8.64
Practical Pistol Course	76.84	9.99
Classroom Average	82.03	5.32
Job Knowledge	6.60	1.67
Judgment	6.41	1.85
Initiative	6.44	1.98
Dependability	6.49	2.14
Demeanor	6.57	1.91
Attitude	6.50	1.88
Relations with Others	6.83	1.69
Communication	6.79	1.87

Primary Manipulations and Analyses

Analysis I

The first *a priori* multivariate factorial analysis tested the effects of the thirteen training school variables with length of academy training treated as a covariate. Race, sex, and mental ability were used as independent variables.

The overall second-order interaction effect of race by sex by mental ability (R×S×M) was nonsignificant at the probability level less than .400 (Table 8). All univariate *F* tests for the interaction were also nonsignificant.

All overall first-order interaction effects (i.e., race by sex (R×S), race by mental ability (R×M), and sex by mental ability (S×M) were nonsignificant. The R×S interaction effect was nonsignificant at the probability level less than .574 (Table 9); the R×M interaction effect was nonsignificant at the probability level less than .572 (Table 10); and the S×M interaction effect was nonsignificant at the probability level less than .830 (Table 11). All univariate *F* tests for each of the interactions were also nonsignificant.

The overall tests of significance for the race, sex, and mental ability main effects indicated that all three factors were significant at the probability level less than .001 (race, Table 12; sex, Table 13; and mental ability, Table 14).

The race effect indicated that there was a significant difference in performance between the black patrol officer group and the white patrol officer group for police academy

Table 8. Multivariate and Univariate Tests of Significance for the Race x Sex x Mental Ability Interaction Effect for Police Academy Training Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	13; 206	1.055	.400
Accident Investigation	Univariate F Test	1; 218	5.150	.024
Traffic Ordinances	Univariate F Test	1; 218	.003	.960
State Traffic Laws	Univariate F Test	1; 218	.160	.682
Speed and Skidmarks	Univariate F Test	1; 218	.546	.461
City Ordinances	Univariate F Test	1; 218	.159	.690
Evidence	Univariate F Test	1; 218	1.119	.291
State Laws	Univariate F Test	1; 218	1.004	.318
Report Writing	Univariate F Test	1; 218	.080	.778
Rules and Regulations	Univariate F Test	1; 218	.610	.436
First Aid	Univariate F Test	1; 218	1.407	.237
Notebook	Univariate F Test	1; 218	.308	.579
Practical Pistol Course	Univariate F Test	1; 218	.950	.331
Classroom Average	Univariate F Test	1; 218	.001	.980

Table 9. Multivariate and Univariate Tests of Significance for the Race x Sex Interaction Effect for Police Academy Training Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	13; 206	.881	.574
Accident Investigation	Univariate F Test	1; 218	.141	.708
Traffic Ordinances	Univariate F Test	1; 218	2.301	.131
State Traffic Laws	Univariate F Test	1; 218	.004	.951
Speed and Skidmarks	Univariate F Test	1; 218	.540	.463
City Ordinances	Univariate F Test	1; 218	2.433	.120
Evidence	Univariate F Test	1; 218	.010	.920
State Laws	Univariate F Test	1; 218	1.899	.170
Report Writing	Univariate F Test	1; 218	.148	.701
Rules and Regulations	Univariate F Test	1; 218	1.588	.213
First Aid	Univariate F Test	1; 218	.003	.958
Notebook	Univariate F Test	1; 218	.403	.526
Practical Pistol Course	Univariate F Test	1; 218	1.808	.180
Classroom Average	Univariate F Test	1; 218	.104	.748

Table 10. Multivariate and Univariate Tests of Significance for the Race x Mental Ability Interaction Effect for Police Academy Training Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	13; 206	.882	.572
Accident Investigation	Univariate F Test	1; 218	.172	.679
Traffic Ordinances	Univariate F Test	1; 218	.058	.810
State Traffic Laws	Univariate F Test	1; 218	.496	.482
Speed and Skidmarks	Univariate F Test	1; 218	.005	.942
City Ordinances	Univariate F Test	1; 218	.008	.929
Evidence	Univariate F Test	1; 218	.209	.648
State Laws	Univariate F Test	1; 218	.001	.969
Report Writing	Univariate F Test	1; 218	.188	.665
Rules and Regulations	Univariate F Test	1; 218	.210	.647
First Aid	Univariate F Test	1; 218	3.489	.063
Notebook	Univariate F Test	1; 218	.215	.643
Practical Pistol Course	Univariate F Test	1; 218	.366	.244
Classroom Average	Univariate F Test	1; 218	.371	.543

Table 11. Multivariate and Univariate Tests of Significance for the Sex x Mental Ability Interaction Effect for Police Academy Training Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	13; 206	,627	.830
Accident Investigation	Univariate F Test	1; 218	.127	.722
Traffic Ordinances	Univariate F Test	1; 218	.240	.625
State Traffic Laws	Univariate F Test	1; 218	1.744	.188
Speed and Skidmarks	Univariate F Test	1; 218	.032	.858
City Ordinances	Univariate F Test	1; 218	.410	.523
Evidence	Univariate F Test	1; 218	.496	.482
State Laws	Univariate F Test	1; 218	.711	.400
Report Writing	Univariate F Test	1; 218	.269	.605
Rules and Regulations	Univariate F Test	1; 218	.423	.516
First Aid	Univariate F Test	1; 218	.027	.871
Notebook	Univariate F Test	1; 218	.424	.516
Practical Pistol Course	Univariate F Test	1; 218	.187	.666
Classroom Average	Univariate F Test	1; 218	.050	.823

Table 12. Multivariate and Univariate Tests of Significance for the Race Main Effect for Police Academy Training Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	13; 206	4.305	.001*
Accident Investigation	Univariate F Test	1; 218	3.817	.052
Traffic Ordinances	Univariate F Test	1; 218	5.992	.015
State Traffic Laws	Univariate F Test	1; 218	1.988	.160
Speed and Skidmarks	Univariate F Test	1; 218	14.599	.001*
City Ordinances	Univariate F Test	1; 218	20.323	.001*
Evidence	Univariate F Test	1; 218	1.021	.313
State Laws	Univariate F Test	1; 218	.034	.854
Report Writing	Univariate F Test	1; 218	5.420	.021
Rules and Regulations	Univariate F Test	1; 218	15.271	.001*
First Aid	Univariate F Test	1; 218	16.482	.001*
Notebook	Univariate F Test	1; 218	3.427	.065
Practical Pistol Course	Univariate F Test	1; 218	1.746	.188
Classroom Average	Univariate F Test	1; 218	19.194	.001*

*Significant variables at established alpha levels

Table 13. Multivariate and Univariate Tests of Significance for the Sex Main Effect for Police Academy Training Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	13; 206	4.805	.001*
Accident Investigation	Univariate F Test	1; 218	10.239	.002*
Traffic Ordinances	Univariate F Test	1; 218	5.266	.023
State Traffic Laws	Univariate F Test	1; 218	11.065	.001*
Speed and Skidmarks	Univariate F Test	1; 218	.026	.872
City Ordinances	Univariate F Test	1; 218	1.321	.252
Evidence	Univariate F Test	1; 218	.025	.874
State Laws	Univariate F Test	1; 218	8.473	.004*
Report Writing	Univariate F Test	1; 218	2.414	.122
Rules and Regulations	Univariate F Test	1; 218	9.113	.003*
First Aid	Univariate F Test	1; 218	8.400	.004*
Notebook	Univariate F Test	1; 218	2.806	.095
Practical Pistol Course	Univariate F Test	1; 218	20.737	.001*
Classroom Average	Univariate F Test	1; 218	12.985	.001*

*Significant variables at established alpha levels

Table 14. Multivariate and Univariate Tests of Significance for the Mental Ability Main Effect for Police Academy Training Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	13; 206	3.451	.001*
Accident Investigation	Univariate F Test	1; 218	2.603	.108
Traffic Ordinances	Univariate F Test	1; 218	5.605	.019
State Traffic Laws	Univariate F Test	1; 218	1.724	.191
Speed and Skidmarks	Univariate F Test	1; 218	11.649	.001*
City Ordinances	Univariate F Test	1; 218	7.912	.005
Evidence	Univariate F Test	1; 218	6.378	.012
State Laws	Univariate F Test	1; 218	15.151	.001*
Report Writing	Univariate F Test	1; 218	7.280	.008
Rules and Regulations	Univariate F Test	1; 218	14.200	.001*
First Aid	Univariate F Test	1; 218	16.813	.001*
Notebook	Univariate F Test	1; 218	.023	.881
Practical Pistol Course	Univariate F Test	1; 218	6.035	.015
Classroom Average	Univariate F Test	1; 218	32.072	.001*

*Significant variables at established alpha levels

training. Examination of the univariate F tests and the corresponding mean values of performance in the training courses for black and white patrol officers indicated that white officers performed significantly better (i.e., received significantly higher scores) on the training dimensions of Speed and Skidmarks, City Ordinances, Rules and Regulations, First Aid, and Classroom Average.

The sex effect indicated that there was a significant difference in performance between the male patrol officer group and the female patrol officer group for police academy training. Examination of the univariate F tests and the corresponding mean values of performance in the training courses for male and female patrol officers indicated that female officers received significantly higher scores on the training dimensions of Accident Investigation, State Traffic Laws, State Laws, Rules and Regulations, First Aid, and Classroom Average. Male officers received significantly higher scores on the training dimension of Practical Pistol Course.

The mental ability effect indicated that there was a significant difference in performance between the higher mental ability group of patrol officers and the lower mental ability group of patrol officers for police academy training. Examination of the univariate F tests and the corresponding mean values of performance in the training courses for higher mental ability and lower mental ability patrol officers indicated that higher mental ability officers received

significantly higher scores on the training dimensions of Speed and Skidmarks, State Laws, Rules and Regulations, First Aid, and Classroom Average.

Since there were no overall significant interactions, the indication was that the main effects are additive. The true model equation for the effects of race, sex, and mental ability using training school dimensions as dependent variables adjusted for length of academy training is:

$$\underline{Y}_{ijk} = \underline{\mu} + \underline{R}_i + \underline{S}_j + \underline{M}_k + \underline{e}$$

The thesis hypotheses relevant to this analysis are: Hypothesis 1: No significant race differences will occur on the police academy training variables on which the subjects are compared. Based upon the significant overall race effect and the five significant univariate F tests for the race effect as depicted in Table 12, this hypothesis is rejected. Hypothesis 2: No significant sex differences will occur on the police academy training variables on which the subjects are compared. Based upon the significant overall sex effect and the seven significant univariate F tests for the sex effect as depicted in Table 13, this hypothesis is rejected. Hypothesis 3: No significant mental ability differences will occur on the police academy training variables on which the subjects are compared. Based upon the significant overall mental ability effect and the five significant univariate F

tests for the mental ability effect as depicted in Table 14, this hypothesis is rejected.

Hypothesis 10: No overall significant interactions will occur on any of the multivariate factorial analyses using the factors of race, sex, and mental ability. Based upon the non-significant overall interactions for this analysis as depicted in Tables 8, 9, 10, and 11, this hypothesis could not be rejected at this point in the data analyses.

Analysis I(a)

Analysis I(a) was conducted after it was determined that significant race and sex differences existed when police academy training dimensions were used as dependent variables. This analysis was performed to determine whether significant race and/or sex differences existed on the mental ability variable as indicated by scores on the Otis-Lennon Mental Ability Test. It was postulated that, if a significant difference was found, mental ability could be considered an important variable contributing to the corresponding race and/or sex differences found on the police academy training variables in Analysis I.

This a posteriori univariate factorial analysis tested the effects of mental ability scores using race and sex as independent variables. Race was the only significant factor obtained in this analysis. The univariate F test for the race effect indicated significance at the probability level less than .001. The white patrol officer group received

significantly higher scores on the Otis-Lennon Mental Ability Test than did the black patrol officer group. The univariate F test for the sex effect was nonsignificant at the probability level less than .172. The univariate F test for the R×S interaction effect was nonsignificant at the probability level less than .612. Therefore, this analysis indicated that mental ability differences between black and white patrol officers contributed to the training performance differences between black and white patrol officers found in Analysis I.

Analysis II

The second a priori multivariate factorial analysis tested the effects of the eight supervisory rating variables with length of academy training treated as a covariate. Race, sex, and mental ability were used as independent variables.

The overall second-order interaction effect (R×S×M) was significant at the probability level less than .041 (Table 15). However, all univariate F tests for the interaction were nonsignificant.

All overall first-order interaction effects (i.e., (R×S), (R×M), and (S×M)) were nonsignificant. The R×S interaction effect was nonsignificant at the probability level less than .522 (Table 16); the R×M interaction effect was nonsignificant at the probability level less than .977 (Table 17); and the S×M interaction effect was nonsignificant at the probability level less than .181 (Table 18). All univariate F tests for each of the interactions were also nonsignificant.

Table 15. Multivariate and Univariate Tests of Significance for the Race x Sex x Mental Ability Interaction Effect for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	2.066	.041*
Job Knowledge	Univariate F Test	1; 218	.389	.533
Judgment	Univariate F Test	1; 218	.952	.330
Initiative	Univariate F Test	1; 218	1.036	.310
Dependability	Univariate F Test	1; 218	.551	.459
Demeanor	Univariate F Test	1; 218	.109	.742
Attitude	Univariate F Test	1; 218	.148	.701
Relations with Others	Univariate F Test	1; 218	.094	.760
Communication	Univariate F Test	1; 218	1.416	.235

*Significant variables at established alpha levels

Table 16. Multivariate and Univariate Tests of Significance for the Race x Sex Interaction Effect for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	.895	.522
Job Knowledge	Univariate F Test	1; 218	.001	.982
Judgment	Univariate F Test	1; 218	.000	.987
Initiative	Univariate F Test	1; 218	.682	.410
Dependability	Univariate F Test	1; 218	.104	.747
Demeanor	Univariate F Test	1; 218	.016	.900
Attitude	Univariate F Test	1; 218	.177	.674
Relations with Others	Univariate F Test	1; 218	.039	.843
Communication	Univariate F Test	1; 218	.204	.652

Table 17. Multivariate and Univariate Tests of Significance for the Race x Mental Ability Interaction Effect for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	.264	.977
Job Knowledge	Univariate F Test	1; 218	.830	.363
Judgment	Univariate F Test	1; 218	.901	.343
Initiative	Univariate F Test	1; 218	.669	.414
Dependability	Univariate F Test	1; 218	.647	.422
Demeanor	Univariate F Test	1; 218	1.381	.241
Attitude	Univariate F Test	1; 218	1.003	.318
Relations with Others	Univariate F Test	1; 218	1.688	.195
Communication	Univariate F Test	1; 218	1.194	.276

Table 18. Multivariate and Univariate Tests of Significance for the Sex x Mental Ability Interaction Effect for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	1.442	.181
Job Knowledge	Univariate F Test	1; 218	.239	.625
Judgment	Univariate F Test	1; 218	.002	.965
Initiative	Univariate F Test	1; 218	1.751	.187
Dependability	Univariate F Test	1; 218	.029	.865
Demeanor	Univariate F Test	1; 218	.473	.492
Attitude	Univariate F Test	1; 218	.225	.636
Relations with Others	Univariate F Test	1; 218	.061	.805
Communication	Univariate F Test	1; 218	.232	.630

The overall tests of significance for the race, sex, and mental ability main effects indicated that race was the only significant factor, being significant at the probability level less than .045 (Table 19). The race effect indicated that there was a significant difference between the black patrol officer group and the white patrol officer group for ratings received on the performance instrument. Examination of the univariate F tests indicated that there were no significant differences on the race factor for any of the performance dimension scales. Although there were so significant differences for any of the rating dimensions taken singularly, as denoted by the univariate F tests, the overall test of significance indicated that the group centroid for the black patrol officer group was significantly different from the group centroid for the white patrol officer group in a multidimensional (i.e., eight dimensions corresponding to the eight rating scales) space. Examination of the mean values of performance evaluation on the rating scales for black and white patrol officers indicated that the white patrol officer group was rated significantly higher (i.e., more favorably) on the supervisory rating scales than was the black patrol officer group.

The sex effect was nonsignificant at the probability level less than .563 (Table 20). This indicated that the male and female patrol officers were not rated significantly different (i.e., neither group was rated significantly higher

Table 19. Multivariate and Univariate Tests of Significance for the Race Main Effect for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	2.028	.045*
Job Knowledge	Univariate F Test	1; 218	1.047	.307
Judgment	Univariate F Test	1; 218	.124	.726
Initiative	Univariate F Test	1; 218	1.803	.181
Dependability	Univariate F Test	1; 218	2.378	.124
Demeanor	Univariate F Test	1; 218	.005	.945
Attitude	Univariate F Test	1; 218	.000	.993
Relations with Others	Univariate F Test	1; 218	.092	.762
Communication	Univariate F Test	1; 218	3.750	.054

*Significant variables at established alpha levels

Table 20. Multivariate and Univariate Tests of Significance for the Sex Main Effect for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	.846	.563
Job Knowledge	Univariate F Test	1; 218	1.537	.216
Judgment	Univariate F Test	1; 218	.390	.533
Initiative	Univariate F Test	1; 218	.523	.470
Dependability	Univariate F Test	1; 218	2.266	.134
Demeanor	Univariate F Test	1; 218	.758	.385
Attitude	Univariate F Test	1; 218	1.698	.194
Relations with Others	Univariate F Test	1; 218	.461	.498
Communication	Univariate F Test	1; 218	.001	.972

nor significantly lower than the other group) on the performance instrument. Examination of the univariate F tests indicated that there were no significant differences on the sex factor for any of the performance dimension scales.

The mental ability effect was nonsignificant at the probability level less than .111 (Table 21). This indicated that the higher and lower mental ability patrol officers were not rated significantly different on the performance instrument. Examination of the univariate F tests indicated that there were no significant differences on the mental ability factor for any of the performance dimension scales.

Since the only overall significant interaction obtained was the second-order interaction effect of R×S×M, this indicated that the simple interaction effects of two factors change as a function of the level of the third factor. Thus, the differences between the groups on one factor are not constant or additive with the differences between the groups on the other two factors. Since there was only a slight significance level for the second-order interaction, and since none of the univariate F tests for the interaction was significant, it was decided to continue analyses of the first-order interaction effects and main effects rather than discontinuing analyses of the supervisory rating variables after finding a significant second-order interaction. An analysis of simple effects was also undertaken to aid the interpretations of the data. Tests of simple effects are essentially tests to

Table 21. Multivariate and Univariate Tests of Significance for the Mental Ability Main Effect for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	1.656	.111
Job Knowledge	Univariate F Test	1; 218	.878	.350
Judgment	Univariate F Test	1; 218	.438	.509
Initiative	Univariate F Test	1; 218	.202	.653
Dependability	Univariate F Test	1; 218	.088	.767
Demeanor	Univariate F Test	1; 218	1.106	.294
Attitude	Univariate F Test	1; 218	.057	.811
Relations with Others	Univariate F Test	1; 218	.439	.508
Communication	Univariate F Test	1; 218	4.738	.031

determine the equality of sets of parameters within levels of other parameters. Appelbaum (1974) has recommended this procedure when a significant interaction occurs in multifactor experimental designs.

Initially, the analysis of simple effects examined the first-order interactions at both levels of the third factor. The overall test of R×S within the higher mental ability level was nonsignificant at the probability level less than .101 (Table 22). The overall test of R×S within the lower mental ability level was nonsignificant at the probability level less than .263 (Table 23). The overall test of R×M within the male level was nonsignificant at the probability level less than .797 (Table 24). The overall test of R×M within the female level was nonsignificant at the probability level less than .088 (Table 25). The overall test of S×M within the black level was nonsignificant at the probability level less than .092 (Table 26). The overall test of S×M within the white level was nonsignificant at the probability level less than .084 (Table 27). All univariate F tests for each of the interactions were also nonsignificant.

An analysis was subsequently performed to test the simple main effects within both levels of the second and third factors. The overall test of race within the male higher mental group was nonsignificant at the probability level less than .745 (Table 28). The overall test of race within the male lower mental ability group was significant at the

Table 22. Multivariate and Univariate Tests of Significance for the Race x Sex Interaction Effect within the Higher Mental Ability Level for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	1.694	.101
Job Knowledge	Univariate F Test	1; 218	.246	.620
Judgment	Univariate F Test	1; 218	.551	.459
Initiative	Univariate F Test	1; 218	1.718	.191
Dependability	Univariate F Test	1; 218	.135	.714
Demeanor	Univariate F Test	1; 218	.112	.738
Attitude	Univariate F Test	1; 218	.318	.573
Relations with Others	Univariate F Test	1; 218	.012	.912
Communication	Univariate F Test	1; 218	.398	.529

Table 23. Multivariate and Univariate Tests of Significance for the Race x Sex Interaction Effect within the Lower Mental Ability Level for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	1.266	.263
Job Knowledge	Univariate F Test	1; 218	.144	.705
Judgment	Univariate F Test	1; 218	.401	.527
Initiative	Univariate F Test	1; 218	.000	.992
Dependability	Univariate F Test	1; 218	.521	.471
Demeanor	Univariate F Test	1; 218	.013	.910
Attitude	Univariate F Test	1; 218	.006	.936
Relations with Others	Univariate F Test	1; 218	.121	.728
Communication	Univariate F Test	1; 218	1.221	.270

Table 24. Multivariate and Univariate Tests of Significance for the Race x Mental Ability Interaction Effect within the Male Level for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	.576	.797
Job Knowledge	Univariate F Test	1; 218	.336	.563
Judgment	Univariate F Test	1; 218	.223	.638
Initiative	Univariate F Test	1; 218	1.348	.247
Dependability	Univariate F Test	1; 218	1.074	.301
Demeanor	Univariate F Test	1; 218	.884	.348
Attitude	Univariate F Test	1; 218	1.148	.285
Relations with Others	Univariate F Test	1; 218	1.721	.191
Communication	Univariate F Test	1; 218	2.196	.140

Table 25. Multivariate and Univariate Tests of Significance for the Race x Mental Ability Interaction Effect within the Female Level for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	1.751	.088
Job Knowledge	Univariate F Test	1; 218	.885	.348
Judgment	Univariate F Test	1; 218	1.632	.203
Initiative	Univariate F Test	1; 218	.356	.552
Dependability	Univariate F Test	1; 218	.123	.726
Demeanor	Univariate F Test	1; 218	.608	.436
Attitude	Univariate F Test	1; 218	.003	.955
Relations with Others	Univariate F Test	1; 218	.062	.804
Communication	Univariate F Test	1; 218	.412	.522

Table 26. Multivariate and Univariate Tests of Significance for the Sex x Mental Ability Interaction Effect within the Black Level for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	1.734	.092
Job Knowledge	Univariate F Test	1; 218	.043	.836
Judgment	Univariate F Test	1; 218	.654	.420
Initiative	Univariate F Test	1; 218	2.582	.110
Dependability	Univariate F Test	1; 218	.486	.487
Demeanor	Univariate F Test	1; 218	.021	.884
Attitude	Univariate F Test	1; 218	.350	.555
Relations with Others	Univariate F Test	1; 218	.155	.695
Communication	Univariate F Test	1; 218	.443	.506

Table 27. Multivariate and Univariate Tests of Significance for the Sex x Mental Ability Interaction Effect within the White Level for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	1.774	.084
Job Knowledge	Univariate F Test	1; 218	.586	.445
Judgment	Univariate F Test	1; 218	.300	.584
Initiative	Univariate F Test	1; 218	.205	.651
Dependability	Univariate F Test	1; 218	.094	.759
Demeanor	Univariate F Test	1; 218	.560	.455
Attitude	Univariate F Test	1; 218	.023	.880
Relations with Others	Univariate F Test	1; 218	.000	.988
Communication	Univariate F Test	1; 218	1.205	.274

Table 28. Multivariate and Univariate Tests of Significance for the Race Effect within the Male Higher Mental Ability Group for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	.638	.745
Job Knowledge	Univariate F Test	1; 218	.004	.952
Judgment	Univariate F Test	1; 218	.045	.833
Initiative	Univariate F Test	1; 218	.185	.667
Dependability	Univariate F Test	1; 218	.000	.989
Demeanor	Univariate F Test	1; 218	.571	.451
Attitude	Univariate F Test	1; 218	.918	.339
Relations with Others	Univariate F Test	1; 218	.777	.379
Communication	Univariate F Test	1; 218	.033	.855

probability level less than .041 (Table 29). The overall test of race within the female higher mental ability group was nonsignificant at the probability level less than .067 (Table 30). The overall test of race within the female lower mental ability group was nonsignificant at the probability level less than .724 (Table 31). All univariate F tests for the simple race effect, including those for the race effect within the male lower mental ability group, were nonsignificant. However, the tests of the simple main effects for race indicated that race was a significant factor differentiating black male patrol officers in the lower mental ability group from white male patrol officers in the lower mental ability group on the performance dimension instrument although there were no significant differences on any of the performance dimension scales. Examination of the mean values of performance evaluation on the rating scales indicated that white male patrol officers in the lower mental ability group were rated significantly higher on the performance instrument than were black male patrol officers in the lower mental ability group.

The overall test of sex within the black higher mental ability group was nonsignificant at the probability level less than .163 (Table 32). The overall test of sex within the black lower mental ability group was nonsignificant at the probability level less than .737 (Table 33). The overall test of sex within the white higher mental ability group was nonsignificant at the probability level less than .344 (Table 34).

Table 29. Multivariate and Univariate Tests of Significance for the Race Effect within the Male Lower Mental Ability Group for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	2.060	.041*
Job Knowledge	Univariate F Test	1; 218	1.358	.245
Judgment	Univariate F Test	1; 218	.296	.587
Initiative	Univariate F Test	1; 218	2.184	.141
Dependability	Univariate F Test	1; 218	3.595	.059
Demeanor	Univariate F Test	1; 218	.316	.574
Attitude	Univariate F Test	1; 218	.240	.625
Relations with Others	Univariate F Test	1; 218	1.112	.293
Communication	Univariate F Test	1; 218	6.042	.015

*Significant variables at established alpha levels

Table 30. Multivariate and Univariate Tests of Significance for the Race Effect within the Female Higher Mental Ability Group for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	1.868	.067
Job Knowledge	Univariate F Test	1; 218	.275	.601
Judgment	Univariate F Test	1; 218	.871	.352
Initiative	Univariate F Test	1; 218	1.561	.213
Dependability	Univariate F Test	1; 218	.162	.688
Demeanor	Univariate F Test	1; 218	.560	.455
Attitude	Univariate F Test	1; 218	.024	.877
Relations with Others	Univariate F Test	1; 218	.099	.754
Communication	Univariate F Test	1; 218	.377	.540

Table 31. Multivariate and Univariate Tests of Significance for the Race Effect within the Female Lower Mental Ability Group for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	.662	.724
Job Knowledge	Univariate F Test	1; 218	.694	.406
Judgment	Univariate F Test	1; 218	.764	.383
Initiative	Univariate F Test	1; 218	.285	.594
Dependability	Univariate F Test	1; 218	.005	.944
Demeanor	Univariate F Test	1; 218	.107	.744
Attitude	Univariate F Test	1; 218	.071	.791
Relations with Others	Univariate F Test	1; 218	.000	.986
Communication	Univariate F Test	1; 218	.074	.786

Table 32. Multivariate and Univariate Tests of Significance for the Sex Effect within the Black Higher Mental Ability Group for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	1.489	.163
Job Knowledge	Univariate F Test	1; 218	.015	.902
Judgment	Univariate F Test	1; 218	.230	.632
Initiative	Univariate F Test	1; 218	2.968	.086
Dependability	Univariate F Test	1; 218	.851	.357
Demeanor	Univariate F Test	1; 218	.051	.821
Attitude	Univariate F Test	1; 218	1.049	.307
Relations with Others	Univariate F Test	1; 218	.183	.669
Communication	Univariate F Test	1; 218	.231	.631

Table 33. Multivariate and Univariate Tests of Significance for the Sex Effect within the Black Lower Mental Ability Group for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	.647	.737
Job Knowledge	Univariate F Test	1; 218	.393	.531
Judgment	Univariate F Test	1; 218	.590	.443
Initiative	Univariate F Test	1; 218	.034	.853
Dependability	Univariate F Test	1; 218	.051	.822
Demeanor	Univariate F Test	1; 218	.011	.917
Attitude	Univariate F Test	1; 218	.374	.542
Relations with Others	Univariate F Test	1; 218	.001	.972
Communication	Univariate F Test	1; 218	.232	.630

Table 34. Multivariate and Univariate Tests of Significance for the Sex Effect within the White Higher Mental Ability Group for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	1.130	.344
Job Knowledge	Univariate F Test	1; 218	1.909	.168
Judgment	Univariate F Test	1; 218	.479	.489
Initiative	Univariate F Test	1; 218	.353	.553
Dependability	Univariate F Test	1; 218	1.162	.282
Demeanor	Univariate F Test	1; 218	1.508	.221
Attitude	Univariate F Test	1; 218	.708	.401
Relations with Others	Univariate F Test	1; 218	.404	.526
Communication	Univariate F Test	1; 218	.190	.664

The overall test of sex within the white lower mental ability group was nonsignificant at the probability level less than .156 (Table 35). All univariate F tests for the simple sex effect were nonsignificant. Therefore, the tests of the simple main effects for sex indicated that sex was a nonsignificant factor within levels of the other two factors (i.e., race and mental ability).

The overall test of mental ability within the black male group was nonsignificant at the probability level less than .563 (Table 36). The overall test of mental ability within the black female group was nonsignificant at the probability level less than .071 (Table 37). The overall test of mental ability within the white male group was nonsignificant at the probability level less than .445 (Table 38). The overall test of mental ability within the white female group was nonsignificant at the probability level less than .078 (Table 39). All univariate F tests for the simple mental ability effect were nonsignificant. Therefore, the tests of the simple main effects for mental ability indicated that mental ability was a nonsignificant factor within levels of the other two factors (i.e., race and sex).

The model equation adopted for the effects of race, sex, and mental ability using supervisory rating dimensions as dependent variables adjusted for length of academy training is:

$$\underline{Y}_{ijk} = \underline{\mu} + \underline{R}_i + \underline{RSM}_{ijk} + \underline{e}$$

Table 35. Multivariate and Univariate Tests of Significance for the Sex Effect within the White Lower Mental Ability Group for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	1.507	.156
Job Knowledge	Univariate F Test	1; 218	.001	.981
Judgment	Univariate F Test	1; 218	.033	.855
Initiative	Univariate F Test	1; 218	.018	.892
Dependability	Univariate F Test	1; 218	1.188	.277
Demeanor	Univariate F Test	1; 218	.004	.953
Attitude	Univariate F Test	1; 218	.151	.698
Relations with Others	Univariate F Test	1; 218	.169	.682
Communication	Univariate F Test	1; 218	1.013	.315

Table 36. Multivariate and Univariate Tests of Significance for the Mental Ability Effect within the Black Male Group for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	.846	.563
Job Knowledge	Univariate F Test	1; 218	.937	.334
Judgment	Univariate F Test	1; 218	.450	.503
Initiative	Univariate F Test	1; 218	1.071	.302
Dependability	Univariate F Test	1; 218	.673	.413
Demeanor	Univariate F Test	1; 218	1.866	.173
Attitude	Univariate F Test	1; 218	.827	.364
Relations with Others	Univariate F Test	1; 218	2.151	.144
Communication	Univariate F Test	1; 218	4.592	.033

Table 37. Multivariate and Univariate Tests of Significance for the Mental Ability Effect within the Black Female Group for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	1.842	.071
Job Knowledge	Univariate F Test	1; 218	.529	.468
Judgment	Univariate F Test	1; 218	1.563	.213
Initiative	Univariate F Test	1; 218	1.626	.204
Dependability	Univariate F Test	1; 218	.132	.717
Demeanor	Univariate F Test	1; 218	.285	.594
Attitude	Univariate F Test	1; 218	.040	.842
Relations with Others	Univariate F Test	1; 218	.095	.759
Communication	Univariate F Test	1; 218	.120	.729

Table 38. Multivariate and Univariate Tests of Significance for the Mental Ability Effect within the White Male Group for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	.990	.445
Job Knowledge	Univariate F Test	1; 218	.421	.517
Judgment	Univariate F Test	1; 218	.085	.771
Initiative	Univariate F Test	1; 218	.278	.599
Dependability	Univariate F Test	1; 218	.464	.497
Demeanor	Univariate F Test	1; 218	.409	.523
Attitude	Univariate F Test	1; 218	.335	.563
Relations with Others	Univariate F Test	1; 218	.000	.996
Communication	Univariate F Test	1; 218	.975	.325

Table 39. Multivariate and Univariate Tests of Significance for the Mental Ability Effect within the White Female Group for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 211	1.802	.078
Job Knowledge	Univariate F Test	1; 218	.356	.551
Judgment	Univariate F Test	1; 218	.232	.630
Initiative	Univariate F Test	1; 218	.411	.522
Dependability	Univariate F Test	1; 218	.011	.915
Demeanor	Univariate F Test	1; 218	.339	.561
Attitude	Univariate F Test	1; 218	.116	.733
Relations with Others	Univariate F Test	1; 218	.000	.989
Communication	Univariate F Test	1; 218	2.140	.145

The thesis hypotheses relevant to this analysis are:

Hypothesis 4: No significant race differences will occur on the supervisory rating variables on which the subjects are compared. Based upon the significant overall race effect as depicted in Table 19, this hypothesis is rejected.

Hypothesis 5: No significant sex differences will occur on the supervisory rating variables on which the subjects are compared. Based upon the nonsignificant differences on the sex factor as depicted in Table 20, this hypothesis is accepted.

Hypothesis 6: No significant mental ability differences will occur on the supervisory rating variables on which the subjects are compared. Based upon the nonsignificant differences on the mental ability factor as depicted in Table 21, this hypothesis is accepted.

Hypothesis 10: No overall significant interactions will occur on any of the multivariate factorial analyses using the factors of race, sex, and mental ability. Based upon the significant overall second-order interaction of R×S×M as depicted in Table 15, this hypothesis is rejected.

Analysis II(a)

Analysis II(a) was conducted after it was determined that an overall significant race difference existed when supervisory rating dimensions were used as dependent variables. This a posteriori multivariate factorial analysis tested the effects of the supervisory rating variables using race of the

supervisor and race of the subordinate as independent variables. The analysis was performed primarily to obtain some indication whether racial bias was operating on the part of the police supervisors who rated the patrol officers. It was postulated that, if a significant interaction of race of the supervisor and race of the subordinate was found, this would provide some evidence that racial bias may have affected the results obtained in Analysis II.

A total of nineteen white supervisors and nine black supervisors rated the patrol officers. Two white lieutenants, one black lieutenant, seventeen white sergeants, and eight black sergeants rated at least one subordinate in the sample. All the supervisors who rated the officers were male. Twenty-five black officers were rated by black supervisors; 38 black officers were rated by white supervisors; 41 white officers were rated by black supervisors; and 123 white officers were rated by white supervisors.

The overall test of significance for the interaction effect of race of the subordinate and race of the supervisor was nonsignificant at the probability level less than .317 (Table 40). None of the univariate F tests for the interaction was significant. Therefore, since there was not an overall significant effect for the interaction nor a significant effect for any of the univariate F tests, the indication was that racial bias did not significantly affect the supervisory ratings.

Table 40. Multivariate and Univariate Tests of Significance for the Race x Race Interaction Effect for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 216	1.172	.317
Job Knowledge	Univariate F Test	1; 223	.109	.742
Judgment	Univariate F Test	1; 223	.137	.712
Initiative	Univariate F Test	1; 223	.009	.923
Dependability	Univariate F Test	1; 223	.006	.940
Demeanor	Univariate F Test	1; 223	.670	.414
Attitude	Univariate F Test	1; 223	.313	.576
Relations with Others	Univariate F Test	1; 223	.022	.882
Communication	Univariate F Test	1; 223	2.031	.156

The overall test of significance for the race effect of the subordinate indicated significance at the probability level less than .009 (Table 41). This indicated that the white patrol officer group was rated significantly higher on the supervisory rating scales than was the black patrol officer group. However, none of the univariate F tests was significant. These results are essentially redundant since the basic effects were found in Analysis II.

The overall test of significance for the race effect of the supervisor was nonsignificant at the probability level less than .136 (Table 42). The only univariate F test which was found significant was for the Job Knowledge variable at the probability level less than .003. Thus, while there was not a significant overall difference between black and white police supervisors in the ratings assigned to the patrol officers, the mean values for the Job Knowledge variable indicated that black supervisors rated the patrol officers significantly lower on this variable than did white supervisors.

Analysis III

The third a priori multivariate factorial analysis tested the effects of the eight supervisory rating variables with length of academy training and the thirteen training school variables treated as covariates. Race, sex, and mental ability were used as independent variables.

The overall second-order interaction effect (R×S×M)

Table 41. Multivariate and Univariate Tests of Significance for the Race Effect of the Subordinate for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 216	2.649	.009*
Job Knowledge	Univariate F Test	1; 223	1.131	.289
Judgment	Univariate F Test	1; 223	.148	.701
Initiative	Univariate F Test	1; 223	1.251	.265
Dependability	Univariate F Test	1; 223	2.265	.134
Demeanor	Univariate F Test	1; 223	.041	.840
Attitude	Univariate F Test	1; 223	.046	.830
Relations with Others	Univariate F Test	1; 223	.144	.705
Communication	Univariate F Test	1; 223	6.546	.011

*Significant variables at established alpha levels

Table 42. Multivariate and Univariate Tests of Significance for the Race Effect of the Supervisor for Supervisory Rating Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 216	1.568	.136
Job Knowledge	Univariate F Test	1; 223	9.216	.003*
Judgment	Univariate F Test	1; 223	2.222	.137
Initiative	Univariate F Test	1; 223	1.909	.168
Dependability	Univariate F Test	1; 223	1.633	.203
Demeanor	Univariate F Test	1; 223	1.859	.174
Attitude	Univariate F Test	1; 223	3.609	.059
Relations with Others	Univariate F Test	1; 223	3.129	.078
Communication	Univariate F Test	1; 223	3.428	.065

*Significant variables at established alpha levels

was nonsignificant at the probability level less than .055 (Table 43). All univariate F tests for the interaction were also nonsignificant.

All overall first-order interaction effects (i.e., (R×S), (R×M), and (S×M)) were nonsignificant. The R×S interaction effect was nonsignificant at the probability level less than .523 (Table 44); the R×M interaction effect was nonsignificant at the probability level less than .972 (Table 45); and the S×M interaction effect was nonsignificant at the probability level less than .098 (Table 46). All univariate F tests for each of the interactions were also nonsignificant.

The overall tests of significance for the race, sex, and mental ability main effects indicated that none of the effects was significant. The race effect was nonsignificant at the probability level less than .076 (Table 47). The univariate F tests indicated that there were no significant differences on the race factor for any of the performance dimension scales. Thus, the black and white patrol officer groups were not rated significantly different on the supervisory rating scales when academy training variables were treated as covariates. The sex effect was nonsignificant at the probability level less than .250 (Table 48). The univariate F tests indicated that there were no significant differences on the sex factor for any of the performance dimension scales. Thus, the male and female patrol officer groups were not rated significantly different on the supervisory rating scales when

Table 43. Multivariate and Univariate Tests of Significance for the Race x Sex x Mental Ability Interaction Effect for Supervisory Rating Variables Adjusted for Police Academy Training Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 198	1.944	.055
Job Knowledge	Univariate F Test	1; 205	.165	.685
Judgment	Univariate F Test	1; 205	.425	.515
Initiative	Univariate F Test	1; 205	1.220	.271
Dependability	Univariate F Test	1; 205	.387	.535
Demeanor	Univariate F Test	1; 205	.251	.617
Attitude	Univariate F Test	1; 205	.126	.722
Relations with Others	Univariate F Test	1; 205	.332	.565
Communication	Univariate F Test	1; 205	1.699	.194

Table 44. Multivariate and Univariate Tests of Significance for the Race x Sex Interaction Effect for Supervisory Rating Variables Adjusted for Police Academy Training Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 198	.893	.523
Job Knowledge	Univariate F Test	1; 205	.004	.950
Judgment	Univariate F Test	1; 205	.001	.973
Initiative	Univariate F Test	1; 205	.475	.492
Dependability	Univariate F Test	1; 205	.253	.616
Demeanor	Univariate F Test	1; 205	.005	.943
Attitude	Univariate F Test	1; 205	.130	.719
Relations with Others	Univariate F Test	1; 205	.112	.739
Communication	Univariate F Test	1; 205	.114	.736

Table 45. Multivariate and Univariate Tests of Significance for the Race x Mental Ability Interaction Effect for Supervisory Rating Variables Adjusted for Police Academy Training Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 198	.280	.972
Job Knowledge	Univariate F Test	1; 205	.926	.337
Judgment	Univariate F Test	1; 205	.565	.453
Initiative	Univariate F Test	1; 205	1.195	.276
Dependability	Univariate F Test	1; 205	1.192	.276
Demeanor	Univariate F Test	1; 205	1.852	.175
Attitude	Univariate F Test	1; 205	1.006	.317
Relations with Others	Univariate F Test	1; 205	1.054	.306
Communication	Univariate F Test	1; 205	1.188	.277

Table 46. Multivariate and Univariate Tests of Significance for the Sex x Mental Ability Interaction Effect for Supervisory Rating Variables Adjusted for Police Academy Training Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 198	1.710	.098
Job Knowledge	Univariate F Test	1; 205	.353	.553
Judgment	Univariate F Test	1; 205	.001	.979
Initiative	Univariate F Test	1; 205	2.097	.149
Dependability	Univariate F Test	1; 205	.028	.868
Demeanor	Univariate F Test	1; 205	.757	.385
Attitude	Univariate F Test	1; 205	.263	.609
Relations with Others	Univariate F Test	1; 205	.175	.676
Communication	Univariate F Test	1; 205	.140	.709

Table 47. Multivariate and Univariate Tests of Significance for the Race Main Effect for Supervisory Rating Variables Adjusted for Police Academy Training Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 198	1.814	.076
Job Knowledge	Univariate F Test	1; 205	.372	.543
Judgment	Univariate F Test	1; 205	.195	.659
Initiative	Univariate F Test	1; 205	.432	.512
Dependability	Univariate F Test	1; 205	.801	.372
Demeanor	Univariate F Test	1; 205	.085	.771
Attitude	Univariate F Test	1; 205	.068	.794
Relations with Others	Univariate F Test	1; 205	.012	.913
Communication	Univariate F Test	1; 205	3.167	.077

Table 48. Multivariate and Univariate Tests of Significance for the Sex Main Effect for Supervisory Rating Variables Adjusted for Police Academy Training Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 198	1.290	.250
Job Knowledge	Univariate F Test	1; 205	2.460	.118
Judgment	Univariate F Test	1; 205	1.132	.289
Initiative	Univariate F Test	1; 205	2.245	.136
Dependability	Univariate F Test	1; 205	5.429	.021
Demeanor	Univariate F Test	1; 205	1.323	.251
Attitude	Univariate F Test	1; 205	3.146	.078
Relations with Others	Univariate F Test	1; 205	2.544	.112
Communication	Univariate F Test	1; 205	.080	.778

academy training variables were treated as covariates. The mental ability effect was nonsignificant at the probability level less than .203 (Table 49). The univariate F tests indicated that there were no significant differences on the mental ability factor for any of the performance dimension scales. Thus, the higher and lower mental ability patrol officer groups were not rated significantly different on the supervisory rating scales when academy training variables were treated as covariates.

Since there were no overall significant interactions, the indication was that the main effects are additive. The model equation adopted for the effects of race, sex, and mental ability using supervisory rating dimensions as dependent variables adjusted for academy training variables is:

$$\underline{Y}_{ijk} = \underline{\mu} + e$$

The thesis hypotheses relevant to this analysis are:
Hypothesis 7: No significant race differences will occur on the supervisory rating variables when police academy training variables are treated as covariates. Based upon the nonsignificant differences on the race factor as depicted in Table 47, this hypothesis is accepted.
Hypothesis 8: No significant sex differences will occur on the supervisory rating variables when police academy training variables are treated as covariates. Based upon the non-

Table 49. Multivariate and Univariate Tests of Significance for the Mental Ability Main Effect for Supervisory Rating Variables Adjusted for Police Academy Training Variables

VARIABLE	TEST	DEGREES OF FREEDOM	F	PROBABILITY LESS THAN
Overall Effect	Multivariate Lambda	8; 198	1.388	.203
Job Knowledge	Univariate F Test	1; 205	.000	.999
Judgment	Univariate F Test	1; 205	.086	.769
Initiative	Univariate F Test	1; 205	2.131	.146
Dependability	Univariate F Test	1; 205	1.146	.286
Demeanor	Univariate F Test	1; 205	.031	.860
Attitude	Univariate F Test	1; 205	.745	.389
Relations with Others	Univariate F Test	1; 205	.058	.810
Communication	Univariate F Test	1; 205	1.599	.207

significant differences on the sex factor as depicted in Table 48, this hypothesis is accepted.

Hypothesis 9: No significant mental ability differences will occur on the supervisory rating variables when police academy training variables are treated as covariates. Based upon the nonsignificant differences on the mental ability factor as depicted in Table 49, this hypothesis is accepted.

Analysis IV

The fourth analysis used Fisher's r to Z transformation, critical values of an approximation to Student's t statistic, and the multistage Bonferroni procedure to determine the significant correlation coefficients in the intercorrelation matrix of the twenty-two dependent variables (i.e., Otis-Lennon Mental Ability Test scores, thirteen training school subject areas, and eight supervisory rating scales) for the total pooled sample of patrol officers.

The intercorrelation matrix is depicted in Table 50. The critical correlation coefficient at the .05 level of significance was calculated to be .24. The Otis-Lennon Test variable correlated significantly with nine of the thirteen police academy training variables, but did not significantly correlate with any of the supervisory rating variables. Forty of the seventy-eight correlation coefficients derived from intercorrelating the police academy training variables were significant. All twenty-eight correlation coefficients derived

Table 50. Intercorrelation Matrix for the Total Pooled Sample of Patrol Officers

		OL	AI	TO	STL	S&S	CO	E	SL	RW	R&R	FA	N	PPC	CA	JK	Jud	I	Dep	Dem	A	R&O
Police Academy Dimensions	Otis-Lennon																					
	Accident Investigation	.23																				
	Traffic Ordinances	.26*	.19																			
	State Traffic Laws	.20	.18	.29*																		
	Speed and Skidmarks	.40*	.20	.24*	.12																	
	City Ordinances	.35*	.24*	.33*	.16	.36*																
	Evidence	.28*	.13	.24*	.15	.30*	.30*															
	State Laws	.28*	.30*	.41*	.34*	.20	.23	.24*														
	Report Writing	.31*	.12	.32*	.25*	.36*	.30*	.12	.50*													
	Rules and Regulations	.40*	.23	.08	.20	.22	.24*	.16	.18	.16												
	First Aid	.36*	.27*	.32*	.30*	.36*	.38*	.27*	.37*	.29*	.31*											
	Notebook	.09	.09	.29*	.07	.17	.23	.17	.21	.25*	-.09	.06										
	Practical Pistol Course	.18	.10	.09	.06	.18	.04	-.07	.22	.08	-.05	.09	.02									
	Classroom Average	.55*	.46*	.57*	.43*	.58*	.60*	.53*	.66*	.64*	.40*	.57*	.36*	.16								
Supervisory Dimensions	Job Knowledge	.10	.02	.08	.02	.10	.12	.23	.02	-.07	.14	.13	-.01	-.04	.12							
	Judgment	.06	.00	.04	-.01	.12	.09	.17	.00	-.04	.17	.15	-.02	-.08	.09	.81*						
	Initiative	.00	.08	.10	.09	.09	.14	.16	.06	.02	.10	.09	.01	-.03	.16	.72*	.75*					
	Dependability	-.01	.09	.12	.09	.03	.13	.14	.05	.03	.11	.07	.05	.00	.15	.69*	.71*	.87*				
	Demeanor	.02	.09	.09	.02	.03	.08	.13	.09	.00	.09	.06	-.03	.04	.12	.66*	.69*	.71*	.79*			
	Attitude	-.04	.06	.06	-.01	.03	.05	.14	.04	-.01	.07	.04	.03	-.04	.08	.74*	.77*	.82*	.83*	.82*		
	Relations with Others	.00	.04	.09	.00	.11	.12	.15	.11	.04	.05	.15	.09	-.08	.15	.73*	.75*	.71*	.70*	.68*	.77*	
	Communication	.18	.09	.13	.04	.11	.12	.20	.13	.02	.13	.17	-.01	.04	.19	.73*	.65*	.64*	.60*	.66*	.62*	.64

from intercorrelating the supervisory rating variables were significant. None of the police academy training variables correlated significantly with any of the supervisory rating variables.

Analysis V

The fifth analysis tested the relationship between the set of police academy training variables and the set of supervisory rating variables for the total pooled sample of patrol officers using canonical correlation. The results of this analysis are depicted in Table 51. The analysis revealed no significant canonical variables at the .05 alpha level. The maximum canonical correlation was obtained in the first weighting of the training variables and rating variables. This canonical correlation of .39 showed a significance level of .184. Therefore, this analysis indicated that there was not a significant relationship between the two sets of variables using scores obtained for the total sample.

The thesis hypothesis relevant to this analysis is: Hypothesis 11: The set of police academy training variables will be significantly related to the set of supervisory rating variables for the total sample. Based upon the nonsignificant canonical variables for the two sets of variables as depicted in Table 51, this hypothesis is rejected.

Table 51. Canonical Correlation Tests of Significance for the Set of Police Academy Training Variables and the Set of Supervisory Rating Variables

Number	Eigenvalue	Canonical Correlation	Wilks' Lambda	Chi-Square	Degrees of Freedom	Significance
1	.15476	.39339	.58222	116.83695	104	.184
2	.12448	.35281	.68882	80.52043	84	.592
3	.09384	.30634	.78675	51.80661	66	.897
4	.05687	.23847	.86823	30.52164	50	.986
5	.04232	.20572	.92057	17.87557	36	.995
6	.02799	.16730	.96126	8.53531	24	.998
7	.00813	.09018	.98893	2.40350	14	1.000
8	.00296	.05438	.99704	.63975	6	.996

CHAPTER IV

DISCUSSION

Based upon the descriptive statistics, the mean raw score on the Otis-Lennon Mental Ability Test for the total sample of patrol officers was generally slightly higher than the mean raw scores reported by Otis and Lennon (1967) for standardization samples of twelfth-grade students and seventeen year-old students. The mean raw scores on the test for white patrol officers were somewhat higher than the mean raw scores for the standardization samples, while the mean raw scores on the test for black patrol officers were somewhat lower than the mean raw scores for the standardization samples. Overall, the average mental ability level of the sample was roughly equivalent to the average mental ability of seventeen year-old, twelfth grade students.

For the total sample, performance on the training variables generally denoted mean grades in the 80's. Performance on the rating scales indicated mean scores between 6.41 (Judgment) and 6.83 (Relations with Others). On each rating scale, the numerical indicators of performance ranged from "1" ("Low") performance level to "9" ("High") performance level, with "5" indicating an "Average" performance level. Therefore, the mean scores for the total sample disclosed somewhat better than "average" behavioral levels for each

performance dimension scale.

The most substantial evidence of subgroup differences was found for the police academy training variables in Analysis I. The race, sex, and mental ability factors showed highly significant differences between their levels. Overall, the white patrol officer group performed significantly better in training than did the black group. The differences were particularly evident on the training dimensions of Speed and Skidmarks, City Ordinances, Rules and Regulations, First Aid, and Classroom Average.

The results of Analysis I(a) revealed that white patrol officers had a significantly higher level of mental ability than black patrol officers. This finding is in accordance with other research which indicates that whites generally score higher on intelligence and mental ability tests than do blacks (Tyler, 1974; Jensen, 1970; Shuey, 1966). Several investigators (e.g., Anastasi, 1976; Brown, 1976; Sax, 1974; Kennedy, 1971; Thorndike and Hagen, 1969) have noted that academic achievement is positively correlated with intelligence and mental ability, that is, persons at higher intelligence levels generally perform better in educational settings than do persons at lower intelligence levels. Therefore, it appears that mental ability contributed substantially to the race differences found in police academy training.

The actual value of formal education as it relates to police performance is an important issue. Cascio and Real

(1976), in a study with the Landy-Farr supervisory rating instrument, disclosed that overall rated performance was significantly related to amount of formal education. Although it was not feasible to obtain data on the educational level of the patrol officers at the time they enrolled in the police academy training program, it is probable that white patrol officers were at a higher educational level than black patrol officers. Based upon national studies, whites generally have a higher level of education than blacks (Hanushek, 1972; Milner, 1972; U.S. Bureau of the Census, 1971; Baughman and Dahlstrom, 1968). Cronbach (1963) has noted that level of education tends to be highly related to intelligence and mental ability. If white patrol officers in the sample were at a higher educational level than the black patrol officers, then it is likely that education, in addition to mental ability, contributed to the race differences. The cultural, socioeconomic, and educational disparities of black people in terms of white values and norms seem to explain the differences between the races on intelligence and in academic settings (Tyler, 1963, 1965, 1974; Bodmer, 1972; Ryan, 1972; Baehr, et al., 1971; Baehr, Furcon, and Froemel, 1968; Lockwood, 1966; Mayfield, 1964; Klineberg, 1963).

Overall, the female patrol officer group performed significantly better in training than did the male group. While male patrol officers received significantly higher scores for the Practical Pistol Course, female officers received

significantly higher scores for Accident Investigation, State Traffic Laws, State Laws, Rules and Regulations, First Aid, and Classroom Average.

Although the results of Analysis I(a) disclosed that there was not a significant difference on mental ability for the male and female patrol officers, the mean scores on the Otis-Lennon Mental Ability Test for the subgroups indicated that females scored somewhat higher than did the males. Thus it seems that mental ability may contribute to differences between males and females in police academy training.

Since intelligence is composed of diverse abilities, it may be postulated that both sexes excel in different areas. Studies have shown that, on the average, males have higher aptitude in visual-spatial relations and orientation, numerical reasoning, mechanical comprehension, and speed and motor coordination, while females have higher aptitude in perceptual speed and accuracy, numerical computation, verbal fluency, rote memory, and manual dexterity (Lindzey, Hall, and Thompson, 1975; Vinacke, 1968; Fryer, Henry, and Sparks, 1965). Since the vast majority of training school courses require high aptitude for verbal materials and memory for high grades in these courses, it is likely that sex differences in training may be attributed to special aptitude requirements of the courses. The aptitude requirements of the Practical Pistol Course may explain the finding that males received higher grades on this training dimension than did

females. The course concludes with a test which is basically physical involving strength, speed, coordination, and motor skill. It is highly likely that the male officers also had more experience in using firearms than did the female officers.

Another potential component which may have contributed to the finding that females performed significantly better overall in training than did males is education. Watson (1967), in a nationwide survey of the educational level of police by sex, found that female officers were significantly better educated than male officers. However, in a more recent study conducted in one police agency (Washington, D.C.), Bloch and Anderson (1974) reported that new policewomen (N of 86) and comparison policemen (N of 86) were similar in education and earned comparable average scores in police academy training. If female patrol officers in the sample were at a higher educational level than the male patrol officers, then it is likely that education contributed to the sex differences.

Finally, it is postulated that achievement motivation may have contributed to the overall sex effect in police training. It is suggested that females who choose the police profession, particularly those who choose it as a long-term career, may represent a diversified sample of the female population. This divergence may be associated with somewhat unconventional values, traits, and interests which may provide an incentive for these women to do well in an occupation traditionally held by males. Raynor (1974) notes that the

relationship between sex role, sex identity, and career striving is a complex issue with regard to explanations for the sex differences found in research on achievement motivation.

Overall, the higher mental ability patrol officer group performed significantly better in training than did the lower mental ability group. The differences were particularly evident on the training dimensions of Speed and Skidmarks, State Laws, Rules and Regulations, First Aid, and Classroom Average. Since academic achievement is highly related to intelligence and mental ability, it is reasonable to assume that mental ability is an important determinant of training performance. Also, it is probable that higher mental ability officers were at a higher educational level than lower mental ability officers since level of education is substantially related to intelligence and mental ability. If, for the sample, the educational level was higher for higher mental ability officers than for lower mental ability officers, then it is likely that education, in addition to mental ability, contributed to the mental ability differences.

Analysis II indicated a significant second-order interaction (R×S×M) and a significant race effect for the supervisory rating variables. Direct interpretation of the race effect for this analysis is difficult due to the overall significant second-order interaction of race by sex by mental ability, and due to the lack of significance for any of the

univariate F tests on the race factor. While the second-order interaction showed a slight level of significance, none of the univariate F tests for the interaction was significant. This indicated that although there might be a slight confounding in a straightforward explanation of the race effect, a general elucidation was justified. Tests of simple effects were undertaken to aid in the interpretation of the results. While the univariate F tests for the race factor failed to reveal specific rating scales which may have accounted for the significant overall indication that differences did exist for black and white patrol officers on the performance scales, the probability levels of significance for the univariate F tests did disclose tendencies among the individual performance scales which contributed to the overall race effect.

Overall, the job performance of white patrol officers was significantly better than the job performance of black patrol officers, as determined by supervisory ratings on eight performance dimensions, although the significance level was slight. The performance scale which appeared to be contributing most to this overall effect was Communication. Other performance scales which seemed to be contributing, but to a lesser degree, to the overall effect were Dependability and Initiative. Analysis of simple effects revealed that the differences between black and white officers were greatest for male officers in the lower mental ability group.

Since the supervisors who rated the patrol officers

in the sample were predominantly white (19 white supervisors and 9 black supervisors rated the patrol officers), this suggested the possibility that the race differences obtained on the supervisory ratings could have been due to prejudicial ratings by white supervisors for black patrol officers. However, results of Analysis II(a) did not indicate the existence of racial bias to a degree that would have significantly affected the results found in Analysis II. Therefore, it appears that the differences between black and white patrol officers on the performance scales are generally representative of actual job performance levels.

If one can assume that training performance transfers to job performance, then one would expect the white patrol officer group to receive significantly higher supervisory ratings than the black group, based on the Analysis I findings. However, this assumption would lead to the expectation that the female patrol officer group would receive significantly higher ratings than the male group, and that the higher mental ability patrol officer group would receive significantly higher ratings than the lower mental ability group. While the white patrol officer group received significantly higher ratings than the black group, neither the female patrol officer group nor the higher mental ability patrol officer group received significantly higher ratings than the male group and the lower mental ability group, respectively. However, there is a possibility that sexual bias against females may have influenced

the performance ratings, particularly since all supervisors who completed rating forms were male. The female group actually received lower ratings than the male group though the differences were not significant. While Bloch and Anderson (1974), in a study with the Washington, D.C. police force, found no evidence of sexual bias on the ratings given male and female patrol officers by 23 trained observers (approximately half being female), Milton, et al. (1974) state that police administrators should anticipate sexual bias as a potential problem in formal evaluation procedures. With regard to the mental ability groups, the higher mental ability group received higher ratings than the lower mental ability group, but the differences were not statistically significant.

Analysis III failed to reveal any significant effects using supervisory ratings as dependent variables adjusted for academy training variables. This analysis was performed to determine whether race differences obtained on job performance, as indicated by ratings on the performance scales, could be attributed significantly to race differences obtained on police academy training performance. Since it was found that a significant race effect existed on the supervisory rating variables adjusted for differences among the subjects on length of academy training (Analysis II), but that the race effect was not significant for the supervisory rating variables adjusted for differences among the subjects on length of academy training and scores received on the thirteen training

school subject areas (Analysis III), it appears that training school performance differences between black and white patrol officers are important components contributing to differences found on patrol performance ratings by supervisors.

Analysis IV was conducted to investigate the relationships among the dependent variables used in the previous analyses. In addition to using training school scores and supervisory ratings as dependent variables, the Otis-Lennon Mental Ability Test variable was converted from a dichotomous independent variable to a continuous dependent variable. It was possible in this analysis to determine the degree to which each of the dependent variables related to all other dependent variables for the total pooled sample of patrol officers.

The most conspicuous property of the intercorrelation matrix (Table 50) is the high positive correlations among the supervisory rating variables. Since all correlation coefficients among these eight variables were significant, this is indicative of a large "halo effect." Thus, if a patrol officer received a high, average, or low rating on one scale of the supervisory rating instrument, then there was a high likelihood that he/she would receive a corresponding high, average, or low rating on the other scales of the instrument.

Examination of the means for the supervisory rating variables generally indicated that the ratings were not unduly stringent or lenient although differences did exist for subgroups. Examination of the standard deviations for the

supervisory rating variables generally indicated that there was sufficient variation in the ratings to remove the possibility of substantial central tendency error. The connotation was that, overall, the raters were not unusually harsh or unusually easy in their ratings, and that there was not an unwillingness on the part of the raters to assign extreme ratings. Generally, supervisors were able to discriminate good performance from poor performance, or at least, were able to express opinions as to good performance versus poor performance.

The correlation coefficients obtained among the training school variables were rather moderate although the majority of the coefficients were significant. This finding disclosed that academy performance for the sample tended to be relatively consistent across school subject areas. The training variable which showed the largest relationship with other training variables was Classroom Average. Generally, Classroom Average can be considered a training performance norm. This variable represents an overall estimate of training performance based on a composite obtained by averaging nine of the other twelve training school variables. The composite score, when related to variable scores upon which it is based, will result in inflated correlation coefficients. The training variable which showed the smallest relationship with other training variables was Practical Pistol Course, which was incidentally the only training school subject area

that measured skills not acquired in an educational setting.

None of the correlation coefficients obtained between the training dimensions and the supervisory rating dimensions was statistically significant. The highest coefficient attained was the relationship between Evidence and Job Knowledge, but the coefficient barely missed significance. A small number of negative coefficients were obtained, indicating that, for those particular variables, subjects who scored somewhat high on one dimension tended to score somewhat low on the other.

The Otis-Lennon Mental Ability Test variable signified a rather moderate relationship with training school variables, and little relationship with supervisory rating variables. It should be expected that the test would relate more to training dimensions than to supervisory rating dimensions since it is basically a test of verbal-educational general mental ability.

Analysis V was conducted to further investigate the relationship between the set of training school scores and the set of supervisory ratings for the total pooled sample of patrol officers. The test of significance for the canonical correlation indicated that there was not a significant relationship between the two sets of variables. The presumption that training performance, as denoted by scores in training school courses or subject areas, transferred to patrol performance, as denoted by ratings on job-related traits or

qualities, was not upheld.

With regard to this lack of a significant relationship between the two sets of variables, it could be speculated that the material acquired or learned in training may be forgotten, that the skills and abilities acquired may degenerate, and that learning how to be a police officer and actually putting the knowledge, skills, and abilities to use in the field setting do not necessarily follow. The problem may be associated with motivational aspects. Another possible explanation for the findings is that both sets of measures attempt to evaluate different types of performance dimensions. While the training school courses are aimed at determining competency in specific tasks and knowledge of specific regulations, the supervisory rating instrument is designed to assess constructs which are intangible to some degree. Training performance appears to be independent of certain qualities and traits deemed important for patrol work. It could be postulated that the qualities and traits associated with the rating instrument may have been altered for many patrol officers from the time they went through training to the time at which the ratings were completed. Such temporal factors would likely affect the degree of relationship between the two sets of variables.

The perspective of training performance as related to job performance was summarized recently in a Brief filed by the Executive Committee of the Division of Industrial-Organizational Psychology (Division 14) of the American

Psychological Association (1975). It stated:

> . . .it is not unusual to be unable to find a statistical, correlation relationship of training performance to subsequent job performance, even though the training program is appropriate in light of job analysis and proper measures of training performance have been devised. Presumably, persons employed for a particular job must first have been trained for it and have demonstrated minimal competence in the training prior to being assigned to the job. Having obtained the necessary knowledge and skills to perform the job, subsequent job performance is likely to reflect other characteristics (e.g., motivation) which are not ordinarily incorporated in the training nor necessarily reflected in training performance. Furthermore, job performance reflects the impact of various factors in the work environment, such as the nature of assignments and relationships with supervisors, which vary for individual employees and may be beyond the control of the employee to influence. . .The methods for measuring job performance also play a critical role in determining whether an empirical relationship between measures of training performance and measures of job performance is meaningful.

CHAPTER V

CONCLUSIONS

This study has explored relationships among performance measures in an attempt to provide some estimate of the validity of a selection instrument (i.e., mental ability test), the effectiveness of the police academy training program, and the job performance of patrol officers. Specific demographic variables (i.e., race, sex, and mental ability) were incorporated in comparative analyses of the sample to determine the extent of subgroup differences. The psychological significance of the study is represented by implications of the findings regarding the importance of demographic variables in moderating performance, the effectiveness of training, and the job performance characteristics of the sample. The improved psychometric value of using a behaviorally-based rating instrument as compared to the more traditional rating forms used in measuring job performance provides a more scientific approach to validating selection and training programs, as well as measuring job performance more accurately. The practical constraints encountered in conducting the study limit the generalizability and conclusiveness of the findings. However, the results embody substantial implications. Discussion has focused on numerous possibilities which may have accounted for or contributed to the findings obtained.

The average patrol officer in the sample is comparable to the average twelfth-grade student on mental ability. The mental ability level is significantly higher for whites than for blacks. Based on performance norms established for the rating instrument, all subgroups were rated somewhat better than average on all performance scales.

Mental ability is a significant factor which contributes to training performance differences. The mental ability selection test is a highly valid predictor of training performance when used as a dichotomous independent variable, but relates only moderately to training performance when used as a continuous dependent variable. The test appears to be invalid as a predictor of patrol performance, as measured by the supervisory rating instrument, when scores for those police applicants who passed the test are related to job performance. Since a restricted range of subjects was used (i.e., subjects who passed the test), this result does not indicate the actual validity of the test, but rather serves as an indirect measure of its utility in the prediction of training performance and subsequent job performance for a selective sample of subjects. It is generally acknowledged that a minimum level of mental ability is necessary to adequately perform the work of a police officer (Spencer and Nichols, 1971; Taylor, 1969; Frost, 1955; Terman, 1917). However, the findings of this study suggest that performance distinctions among the select sample of police officers depend

less on mental ability and more on other factors (e.g. demographic, organizational, psychological, situational).

The conclusion which resulted from use of demographic variables is that race, sex, and mental ability are more important in moderating police academy training proficiency than in moderating actual patrol performance. While race, sex, and mental ability were significant factors for training proficiency, race was the only factor found to be significant for both training proficiency and patrol performance.

Performance ratings on the supervisory scales were not predicted by Otis-Lennon Mental Ability Test scores, nor were they predicted by grades on the training school variables. For the total pooled sample of patrol officers, the performance description scales have little practical validity with these predictor measures. The supervisory rating instrument appears to be more valid if race is used as a moderator variable than if sex and mental ability are used as moderator variables. With regard to mental ability, the lack of a significant difference on rating dimensions confirms the postulation that mental ability becomes less significant a factor in differentiating a select group of patrol officers who have passed numerous selection and training tests than it would be for differentiating a group of applicants in an initial screening.

The differences found between black and white officers on the supervisory rating variables may be attributable to

differences in mental ability and training proficiency. The sex differences found on police academy training variables can generally be ascribed to special aptitude requirements of the courses where females tend to show superiority. The single training dimension on which males showed superiority--the Practical Pistol Course--can be explained by male dominance in the aptitudes required by the test.

The police academy training variables tended to be moderately related. Subject areas in the police academy training program do not appear to have a high interdependency, but do appear to be moderately consistent since many of the courses require memory and verbal skills. There was a sizable halo effect with the supervisory rating variables, and therefore, a strong relationship existed among the performance scales. While other distributional errors (i.e., leniency, severity, central tendency) were not substantial, the amount of error associated with the halo effect could not be determined. The only conclusion which could be drawn from the relationships among the rating scales was that the individual patrol officer tended to receive similar ratings on the scales.

The most conclusive finding of this study is the absence of a significant relationship between the set of training school scores and the set of supervisory ratings of patrol performance for the total pooled sample of patrol officers. In effect, training proficiency, as measured by

police academy training scores, does not seem to be a valid predictor of patrol performance, as measured by supervisory ratings. Patrol performance appears to be influenced by numerous variables not embodied in the training program nor reflected in training performance.

Performance evaluation is apparently more systematic in academy training than in patrol work. The factors which contribute to this discrepancy are: (a) The academy training program formally requires a thorough assessment of police recruits in specific subject areas, whereas the evaluation of patrol performance tends to be informal, rudimentary, and incomplete; (b) Academy training is conducted in a controlled environment with specific courses, tests, and objectives in assessing police recruits, whereas the job environment of a patrol officer is less organized and controlled. The patrol officer is generally not assigned specific, predetermined tasks, and is relatively independent in the manner he/she uses patrol time. The supervision of police recruits in training is generally more structured and extensive than the supervision of patrol officers; (c) The performance criterion problem is easily solved in training, whereas, in the evaluation of patrol performance, there is no general consensus regarding the exact manner in which it is to be assessed; and (d) There appears to be a general unwillingness to assess patrol performance. This resistance can be traced to lack of management understanding of the operating problems and

standards entailed, lack of supervisory training in performance evaluation, confusion regarding performance aspects to be assessed, confusion regarding the purpose to which these evaluations will be used, and skepticism regarding the usefulness or value of performance evaluation.

As is true of many field studies conducted under practical constraints, limitations of the study exist. Caution should therefore be used in the interpretation of the results, especially due to the small numbers of minority officers, particularly female patrol officers, in the sample. However, the results of this study may be generalized to other police agencies provided that numerous situational variables are controlled or resolved in some way to increase the similarities among samples. Some of the situational variables which affect the generalizability of research findings include demographic characteristics (e.g., race, sex, mental ability), selection program characteristics (e.g., types of selection instruments utilized, performance standards for the selection instruments, importance attached to each selection instrument), training program characteristics (e.g., types of training methods utilized, quality of instruction, course content), and organizational characteristics (e.g., administrative policies and procedures, supervisory leadership style, patrol assignments).

CHAPTER VI

RECOMMENDATIONS

This study has yielded several findings which provide suggestions for future research and potential improvement in the areas of recruitment, selection, training, and performance evaluation of patrol officers.

A deficiency of this study was failure to obtain data on educational levels of patrol officers at the time they enrolled in the police academy training program. It is postulated that educational level was an important variable contributing to some of the differences obtained in Analysis I. The educational level of the subjects should be investigated to determine whether significant differences existed with regard to race, sex, and/or mental ability. With regard to the achievement motivation postulate for sex differences in training school proficiency, it is suggested that this hypothesis be tested for females through psychological assessment (e.g. Thematic Apperception Test, Strong Vocational Interest Blank, Kuder Preference Record, Edwards Personal Preference Schedule, Minnesota Multiphasic Personality Inventory). Perhaps an attitude survey could be administered to female patrol officers in the sample to determine the extent of liberated and unconventional attitudes regarding the feminine role. For this

attitude survey, a control group of females from the general population could be used as a comparative base. It is generally acknowledged that an evaluation of mental ability is necessary as an initial step in the selection process. Additionally, numerous researchers have provided evidence that other abilities and traits (e.g., good judgment, common sense, emotional stability, dependability, self-reliance, initiative, personal and moral integrity, courtesy, objectivity, cooperation, sensitivity, accurate observation, ability to factually report observations, stress resistance, self-control, professional demeanor, adaptability, loyalty to the department, alertness, courage) are essential for effective police work. Therefore, in addition to a comprehensive selection program using the traditional measures (e.g., background investigation, medical examination, oral interview, polygraph test, mental ability test), it is recommended that these other essential abilities and traits be measured through various psychological assessment techniques (e.g., California Test of Mental Maturity, Allport-Vernon Study of Values, Gordon Survey of Personal and Interpersonal Values, Guilford-Zimmerman Temperament Survey), situational tests, behavioral observation, self-report tests, and projective tests. If economically feasible, an assessment center could be instituted for police officer selection testing incorporating a comprehensive program designed to evaluate various job-related skills and abilities. The Atlanta Regional Commission (1974) and the Baehr, Furcon,

and Froemel (1968) studies represent research which has undertaken systematic and extensive evaluation of police performance. The former study undertook development of a performance appraisal system from an initial job analysis utilizing the Position Analysis Questionnaire and the Critical Incident Technique. A "Police Officer Performance Evaluation Package" emerged from this comprehensive validation research project. The performance package represents a multidimensional approach to job performance using job-related criterion measures. Such studies are usually quite costly, but the potential benefits generally outweigh the drawbacks.

The recruit training program should be continually assessed. There is a critical need to perform more extensive research on training needs and training objectives, and optimal methods for meeting such needs and objectives. There should be a systematic plan to periodically reevaluate training methods and techniques, course content, instructors, and facilities, and revise or change these training aspects when deemed appropriate. The effectiveness of the academy training program could be increased through a broader scope of training techniques and instructional strategies (e.g., crisis intervention, human relations training, dramatizations, psychodramatic methods, seminars, small-group discussions, interactional training, community relations programs, simulations, interpersonal conflict training, role playing, t-group/

sensitivity training, self-disclosing behavior, feedback sessions, ethnical relations) and multimedia approaches (e.g., lectures, films, videotape, displays, charts, assigned readings). Recruits should be taught to understand the multiplicity of their roles and how to deal with various stressful and unusual situations. It is important that periodic in-service training be provided police officers in order to better prepare them for changes in laws, methodology, operations, and community problems.

Selection and evaluation of field trainers should be performed systematically in order to assure that new officers are receiving proper introduction to the duties and roles of the patrol officer. It has been reported that many police recruits graduate from the police academy with a great deal of motivation, ambition, and idealism, only to be disillusioned by the cynical attitude of the field trainer and veteran officers (Weldy, 1976; Weston, 1976; Bent, 1974; Richardson, 1974; Saunders, 1970; Westley, 1970; McNamara, 1967; Niederhoffer, 1967). The probationary period is a critical period for a new police officer, and the field officers who train and supervise the new officers should provide encouragement, guidance, support, and close supervision to facilitate adequate preparation and evaluation of the police rookie for independent field work (Weldy, 1976; Weston, 1976; Bopp, 1974; Badalamente, et al., 1973; Harris, 1973; McManus, et al., 1970; Wilson, 1962).

A scientific personnel program for a police agency is dependent upon a thorough job analysis to discover the knowledge, skills, and abilities necessary for police work; formulation of selection and training programs based upon job performance requirements; setting of training, selection, and performance objectives; formulation of efficient, effective, and equitable policies and procedures; determination of training needs; and continuous evaluation and revision of selection, training, and performance evaluation, as well as assessment of policies and procedures. Such an analysis might yield results helpful to both the selection and training of police officers.

Since there was a sizable halo effect with the rating scales, it appears that a more valid assessment of police performance depends on the utilization of multiple criteria. Performance evaluation of patrol officers should incorporate both subjective and objective measures emphasizing observable behavior when possible. Since the job of a police officer is multidimensional, single measures alone are generally inadequate in attempting to obtain an accurate evaluation of performance. Both qualitative and quantitative behavioral aspects should be considered as part of an extensive assessment program. The ultimate goal is to obtain a comprehensive assessment of patrol performance by utilizing numerous job-related criteria weighted for importance in a performance index equation. An adequate assessment of patrol performance

must take into account trait dimensions as well as job skill proficiency. Some of the measures which should be considered in constructing an assessment program for patrol officers include supervisory and peer ratings, personal effectiveness measures (e.g., absenteeism, commendations, disciplinary actions, citizen complaints), productivity measures (e.g., number of arrests, number of citations issued, clearance rate, number of convictions of arrestees), training performance, situational tests, job knowledge tests, and tests which allow assessment of skills in particular police tasks.

There have been relatively few validation attempts undertaken with police performance criteria. There is a critical need for more predictive and concurrent validation research utilizing cross-validation and differential validation. A potentially valuable area of research regarding the evaluation of police performance in metropolitan police agencies resides in subgrouping police officers according to some demographic characteristic, such as race or sex, and then comparing subgroup performance levels in an effort to determine the degree of similarity among subgroups. Such comparisons are important for: (A) differential validation; (B) suggesting differential recruitment efforts based on subgroup differences; (c) suggesting modifications in the selection program; (D) suggesting whether supplemental training would be beneficial for police officers in a specific subgroup or subgroups; (E) suggesting characteristics which

may moderate observed performance differences; (F) suggesting job classifications which may be more compatible with the abilities and characteristics of a specific subgroup or subgroups; (G) suggesting possible differential treatment accorded the subgroups by administrative personnel; and (H) suggesting bases on which to make promotional and demotional decisions, transfer decisions, terminations, and salary increments.

APPENDICES

APPENDIX A

TASK ANALYSIS OF THE PATROL SECTION OF THE ATLANTA POLICE DEPARTMENT

In 1974, the City of Atlanta Personnel Office, in conjunction with the Planning and Research Section of the Atlanta Police Department, conducted a task analysis of the Patrol Section of the Police Department. Seventy-two police officers in different ranks and organizational branches of the Police Department were interviewed. Patrol activities and task elements were subsumed under six dimensions. These dimensions are:

(A) Preparation to go on Duty;

(B) Clerical-Type Activities While on Duty;

(C) Actual Police Activities;

(D) Decision-Making;

(E) Preparation to Go Off Duty; and

(F) Additional Responsibilities.

Although a separate task analysis was conducted for each shift (i.e., morning watch, day watch, and evening watch), the activities and responsibilities were virtually identical with slight modifications in the emphasis placed on the duties dependent on watch shift. The more obvious duties of a patrol officer were not included in the original task analysis. However, the experimenter conducted an observational analysis

and discovered other activities and task elements in the work domain of an Atlanta patrol officer. These additional elements (designated by an asterisk) are included in the categories below. Activities and task elements may be identified as follows:

(A) Preparation to Go on Duty

(1) Communicates with officer coming off duty to obtain any necessary information and details including the activity on the beat.

(2) Checks information board for announcements, bulletins, and other job-related information.

(3) Checks information desk for subpoenas, court notices, appointments, court appearances, and appearances before and notices from the Report Review Committee.

(4) Checks for reports on blood tests, breath tests, and other laboratory tests on persons he/she has arrested or detained on suspicion of driving under the influence of alcohol, drug abuse, etc.

(5) Reads daily bulletin and information sheet regarding recent criminal offenses committed within the zone precinct.

(6) Examines "pin maps" for the location of recent robberies, burglaries, and other criminal offenses within the zone precinct.

(7) Obtains subpoenas if needed.

(8) Checks to make sure he/she has general copybook, subpoena book, report forms, citation pad, and radio logbook.

(9) Checks out radio.

*(10) Checks uniform and equipment.

(11) Stands for roll call, listens to briefing from supervisor regarding any pertinent information, and receives daily assignment and/or weekly work sheet from supervisor.

(12) Visually inspects police vehicle before leaving precinct headquarters as a safety precaution.

(13) Transfers personal belongings and work supplies to patrol car.

(B) Clerical-Type Activities While on Duty

(1) Enters radio calls received on daily logsheet.

(2) Enters radio calls investigated on daily logsheet.

(3) Issues traffic citations.

(4) Writes arrest reports, crime-against-person reports, and miscellaneous incident (e.g., accident, injury) reports.

(5) Writes inventory sheet concerning vehicles that are impounded.

(6) Completes juvenile detention forms.

(7) Writes down descriptions of suspects, missing persons, vehicles stolen and missing, and miscellaneous situations supplied by radio contact and individual citizens approaching the officer.

(C) Actual Police Activities

(1) Patrols assigned district maintaining surveillance for possible criminal activities and unruly or unusual situations.

(2) Maintains surveillance of critical areas (e.g., high-crime districts, ghetto areas, parks, parking lots, alleyways, isolated zones) for suspicious persons, and unusual activity within the assigned district.

(3) Maintains surveillance for traffic and vehicular violations (e.g., speeding, illegal parking, expired license, expired inspection sticker, load projection violation).

(4) Maintains surveillance for stolen vehicle, wanted suspect, missing person, etc.

(5) Checks establishments (e.g., commercial, business, church) for security purposes, safety code violations, suspicious persons, unusual activities, unusual situation, and to check with persons operating such establishments.

(6) Performs initial investigation of an establishment (e.g., business, house, apartment) where a criminal activity has occurred.

(7) Calls in descriptions of perpetrators, stolen or missing vehicle, and incidents he/she is investigating.

(8) Listens by radio for activity within the zone precinct.

(9) Acknowledges calls received on the radio.

(10) Answers accident calls, domestic calls, disturbances calls, disorderly conduct calls, burglar alarm calls, etc.

*(11) Responds to and handles emergency calls and situations.

(12) Notifies precinct headquarters of arrival time on radio calls.

(13) Requests calls to THOR (i.e., Target Hardening Opportunity Reduction) Center if burglary has occurred in his/her assigned district.

*(14) Investigates accidents.

*(15) Administers first aid and/or calls for medical or rescue vehicles in emergency situations.

(16) Contacts medical examiner if a person is found dead.

*(17) Intervenes in critical situations (e.g., threatened suicide, family crises, domestic disputes, disturbances involving the mentally ill, alcoholics, and drug addicts), and, where appropriate, directs these individuals to rehabilitative services.

*(18) Provides comfort, support, and guidance to individuals and families involved in tragic circumstances.

(19) Backs up other beat officers if needed in critical situations (e.g., bank robbery, kidnapping, threats

involving explosives).

(20) Disarms dangerous individuals in an appropriate manner.

(21) Calms individuals involved in extraordinarily stressful situations (e.g., serious traffic addident, serious injury)

(22) Regulates and controls traffic when an accident or other irregularity occurs.

(23) Pursues speeders.

*(24) Enforces federal and state laws, and city and county ordinances.

(25) Arrests suspects for alleged law violation.

*(26) Searches suspects if appropriate.

*(27) Reads citizen rights to arrested suspect.

(28) Notifies detectives of an arrest if appropriate.

(29) Checks with appropriate bureau of the police agency for outstanding warrants on suspects and arrested persons.

*(30) Makes necessary reports and records.

(31) Transports arrested persons to detention center.

(32) Exchanges information with other officers at the scene of a crime, accident, or other incident which has occurred.

*(33) Safeguards property.

*(34) Collects, preserves, and safeguards evidence.

(35) Investigates citizen complaints where appropriate.

(36) Interrogates victims, witnesses, and suspects.

(37) Talks with area residents to obtain information concerning any illegal activities within his/her assigned district.

(38) Talks with citizens to establish rapport and improve police-community relations.

(39) Gives information, directions, and advice to citizens who approach him/her.

(40) Mediates and attempts to resolve conflicts and disputes.

(41) Assists in crowd control situations.

*(42) Cooperates with other police agencies and allied units.

*(43) Operates and maintains patrol vehicle and job-related equipment.

(D) Decision-Making

(1) Decides how to handle patrol area (i.e., what activities and problems to devote most attention to, what duties can be combined to cover beat).

*(2) Decides on what approach to take in delicate situations (e.g., suicide threat, domestic disturbance).

*(3) Decides on the probable circumstances surrounding incidents, problem situations, criminal actions, etc.

*(4) Decides on approaches to take in investigating a situation, obtaining information, and securing evidence.

*(5) Decides on whether to detain a suspicious person.

*(6) Decides on whether to arrest a person, issue a warning, issue a citation, etc.

*(7) Decides on whether to refer a citizen to another person, agency, institution, etc.

*(8) Decides on amount of caution and/or amount of force to use in situations.

(9) Determines method of entry into burglarized or vandalized establishment.

(10) Determines approximate value of stolen items.

(11) Determines if drivers are intoxicated by watching them walk 10-15 feet and observing their speech patterns.

(12) Decides on whether to follow-up on any particular incident reported.

(13) Anticipates whether an incident reported will require additional help.

(E) Preparation to Go Off Duty

(1) Returns to precinct headquarters promptly.

(2) Turns in radio.

(3) Turns inventoried items from impounded cars into station supervisor.

(4) Sorts traffic tickets for traffic court, police department, and state patrol.

(5) Turns in copies of citations issued.

(6) Turns in reports.

(7) Fills out daily activity report.

(8) Communicates with officer coming on duty to give any necessary information and details including the activity on the beat.

(9) Turns patrol car over to officer on next watch shift.

(F) Additional Responsibilities

*(1) Meets with other departmental specialists (e.g., detectives, special investigations, vice control, special operations, public information) to discuss relevant situations or circumstantial aspects of an incident.

(2) Testifies in court.

(3) Participates in special training procedures and in-service training.

APPENDIX B

TRAINING SCHOOL SUBJECT AREAS

Accident Investigation. Written examination based on methods, procedures, and instruments for the investigation of traffic accidents, the preparation of accident reports and field notes pertaining to a traffic accident, and the restoration of traffic control at an accident scene. The course places emphasis on approaching an accident scene, questioning witnesses, identifying pertinent physical evidence at a traffic scene, examining vehicles involved in accidents, and checking roadways, signs, and signals.

Traffic Ordinances. Written examination based on knowledge of traffic ordinances of the City of Atlanta. Topics addressed in the course include pedestrian codes, bicycle codes, motorcycle codes, emergency vehicle codes, business district traffic codes, traffic signal codes, towing regulations, impounding regulations, vehicular lighting regulations, parade permit regulations, load projection regulations, speed zone regulations, parking zone regulations, and the issuance of traffic citations. The course is based on Chapter 18 of the City Code.

State Traffic Laws. Written examination based on methods and procedures for the identification and citation of

motor vehicle violations as determined by the Georgia Motor Vehicle Code. The course places emphasis on Georgia motor vehicle laws including various court rulings relating to vehicular codes, license classification, registration and tags, and the Uniform Act regulating traffic. The course is based on Title 68A of the Georgia Code entitled "Uniform Rules of the Road."

Speed and Skidmarks. Practical and written examination based on the determination of minimum speed from the measurement of skidmarks. A skidmark-speed chart utilizing a conversion formula based on skidding distance and skid resistance factors is used to determine speed of a vehicle.

City Ordinances. Written examination based on knowledge of general city ordinances, alcoholic beverage licensing ordinances, and ordinances related to animals. General city ordinances refer to a variety of regulations governing such subject areas as disorderly conduct, false representation of fact, loitering, curfew, noise, safety, malicious mischief and trespass, fireworks, impersonation, summons, the distribution of drugs and medicines for advertisement, interference with a law enforcement officer, emergency powers of the mayor, citizenship and residence requirements, and regulations associated with the unlawful sell, transfer, and possession of property. The course is based on Chapters 5, 7, and 20 of the City Code.

Evidence. Written examination based on an understanding

of the influence of Constitutional Law upon police methods and procedures. The course deals with the organization and presentation of evidence in a court of law. Emphasis is placed on the understanding, identification, and reporting of the elements of a crime necessary for courtroom prosecution. The police recruit is introduced to procedures for responding to direct examination and cross-examination. The course also provides an analysis of the functions of the court, judge, and jury.

State Laws. Written examination based on knowledge of the origin, theory, and present-day application of law including the process of enacting legislation, researching criminal statutes, and interpretation of Georgia law. The course distinguishes among common, statutory, civil, and criminal law. Emphasis is placed on legal codes governing police practices and procedures as they relate to such areas as assault, battery, search, seizure, interrogation, arrest, use of force, jurisdiction, prosecution, frisking, self incrimination, entrapment, eye witness identification, perjury, criminal damage, criminal trespass, obstructing or hindering a law enforcement officer, intent to commit crime, degree of a crime, time and age limits for crimes and criminals, misdemeanors, felonies, electronics surveillance, and affidavits and warrants. The course is based on Title 26 of the Georgia Code.

Report Writing. Written situational examination designed to assess knowledge and effectiveness in preparing

field notes, rough sketches, and incident reports. The course is based on a knowledge and understanding of the Uniform Crime Report Index.

Rules and Regulations. Written examination based on knowledge of 636 rules, regulations, codes, policies, and procedures established for the organization, government, and operation of the Atlanta Police Department. The course is based on a publication of the Atlanta Police Department entitled "Rules and Regulations."

First Aid. Written examination based on the advanced American Red Cross First Aid Course in Emergency Care. Topics dealt with include short distance transfer, emergency rescue, respiratory emergencies, artificial respiration, cardiopulmonary resuscitation, bleeding control, wound dressing and bandaging, bone and joint injuries, poisoning, shock, cold exposure and frostbite, heat exhaustion and heat strokes, burns, swallowed objects, choking, drug abuse, sudden illness, specific injuries, emotional problems, and legal aspects of administering aid.

Notebook. A subjective/objective measure based on the neatness, content, and completeness of the police recruit's notebook. The recruit's notebook is comprised of notes taken in class for the training school curriculum. A weekly score is assigned to each recruit based on the neatness, content, and completeness of notes taken in classes for the previous week. The content requirement is one handwritten page of

notes or one/half typewritten page of notes per hour of class. An average score is computed from the weekly scores assigned to each recruit.

Practical Pistol Course. Practical examination designed by the Federal Bureau of Investigation. The weapon used for the examination is a Smith & Wesson 38-caliber police special. The target is a silhouetted figure of a person on a paper background. After every five rounds of ammunition fired, the recruit must eject the shells and reload with five more rounds. The examination is in two parts. The first part of the test has a 25-second time limit in which the recruit draws his/her weapon, fires five rounds (double action) at the target, ejects the shells, reloads, and fires five more rounds (double action) at the target seven yards away. The second part of the test has a 4-minute and 35-second time limit in which the recruit fires five rounds (single action) from a prone position sixty yards from the target. The recruit then moves ten yards closer to the target and fires five rounds (single action) from a sitting position. Next the recruit fires five rounds (single action) from a prone position fifty yards from the target. Then the recruit moves behind a barricade and fires five rounds (single action) with the preferred hand, and fires five rounds (single action) with the other hand at the target fifty yards away. The recruit then moves twenty-five yards from the target and fires five rounds (double action) from a kneeling position. The last step consists of

moving behind a barricade twenty-five yards from the target and firing five rounds (double action) with the preferred hand, and firing five rounds (double action) with the other hand. The maximum number of rounds fired is 50, with a possible 25 being fired single action and 25 being fired double action. Total score is based on the speed and accuracy of test performance.

Classroom Average. A score based on the average grade for all courses taken in the training school curriculum except the practical pistol course, first aid course, and notebook grade.

APPENDIX C

CONSTRUCTION AND EVALUATION OF THE LANDY-FARR SUPERVISORY RATING INSTRUMENT

Two initial workshop conferences were held in Chicago, Illinois for the purpose of identifying and defining relevant dimensions which exhaustively described the job of police patrol officer. Also, several specific behavioral items were generated at these meetings. Tentative dimensions of patrol performance were identified and defined, and descriptions of high, moderate, and low levels of performance were developed for each dimension. First, the assembled supervisory personnel and police officers were requested to list as many words as possible which described the job of police patrol officers. At the completion of this master list, combinations were effected and duplications were eliminated wherever the supervisory personnel deemed such action appropriate. Second, all officers were required to agree upon a definition of each dimension which survived the first step. Wherever consensus was unattainable, that dimension was deleted from the list. Third, all officers were asked to independently generate several specific behavioral descriptions for each dimension which passed the second step. Only those job dimensions which were common to both supervisory conferences were retained for further scale development. Eight performance dimensions evolved

from the supervisory conferences.

At this point, preparation of the behavioral items (reflecting high, moderate, and low levels of performance) for reallocation to dimensions, and for preliminary scaling procedures was begun. Initial item reallocation studies were undertaken to determine whether items written for a particular performance dimension would be judged as being relevant to that dimension by a majority of supervisors who judged the items. Ambiguous items were subsequently rewritten, and new items were constructed for those dimensions to which less than ten items had been reallocated. Additional behavioral items were obtained by mail from those supervisory police officers who had participated in the two workshop conferences. Lists of all the specific behavioral items generated earlier were sent to police agencies, along with lists of the job dimensions and their definitions. The supervisory police officers were directed to read each specific behavioral item and assign (reallocate) it to whichever dimension they deemed most appropriate. Continued refinement of the developing performance scales occurred as reallocation data were received by the project staff. Upon return of the completed lists, the project staff kept only those items which were consistently assigned (67%) to one job dimension. These items were retained for further scale development.

The project staff visited each police agency which participated in the scaling of specific behavioral items which

had survived the reallocation stage. Supervisory police officers were each given a packet containing all the behavioral examples and the dimensions to which they had been reallocated. The task of each supervisor was to read each specific behavioral item and determine just "how much" of the dimension, to which it had been reallocated, was represented by it. That is, each supervisor was to evaluate each behavioral item and decide whether it was an example of above average, average, or below average behavior with respect to the dimension to which it had been reallocated. The supervisory officers were required to translate their evaluations into ratings conforming to a nine-point scale, where a rating at the top end of the scale indicated a favorable (desirable) behavior and a rating at the lower end of the scale indicated an unfavorable (undesirable) behavior.

Upon completion of the analysis of the scaling data, the project staff selected those behavioral items for inclusion in the performance appraisal scales which conformed to the following criteria: (1) a group of items was selected for each job dimension so that the mean scaling values, as determined from judgments on the nine-point scale, were dispersed as evenly as possible along the range of the entire rating scale; and (2) each selected item had a mean value characterized by a small standard deviation, indicating considerable agreement among the supervisory officers who performed the scaling task regarding the desirability or

undesirability of the behavior described by the item.

Field testing of the supervisory rating instrument consisted of research endeavors designed to obtain psychometric information, determine inter-rater reliability, and determine the validity of the instrument. Psychometric information included measures of central tendency and range, and indices of halo and leniency error. In order to determine inter-rater reliability, it was necessary that at least two supervisors rate the job performance of the same set of patrol officers. Validation studies entailed comparing the ratings given a particular patrol officer with other indices of that patrol officer's value to the police agency. Such indices included other performance ratings recorded on departmental rating scales, training academy performance, disciplinary and commendation records, oral board interview scores, and assessment center ratings. Arrest and violation records were also examined.

Means and standard deviations were determined for 1,469 rating forms. The means (ranging from 6.26 for "Initiative" to 6.56 for "Demeanor") indicated that the ratings were skewed in the direction of high performance. The standard deviations (ranging from 1.65 for "Judgment" to 1.87 for "Initiative") indicated that there was an adequate variation in the ratings for concluding that supervisors were able to discriminate among subordinate patrol officers on each of the performance dimensions.

The median intercorrelation of each of the eight performance scales with the seven other scales ranged from .564 to .705 based on data from thirty-seven cities.

A consideration of the standard deviations indicates that there was sufficient variation in the ratings to rule out the presence of extreme central tendency error on the part of the raters.

Inter-rater reliabilities for the supervisory rating scales were determined from data obtained in eight police agencies. Median reliability coefficients across these eight agencies were obtained for each dimension. The median reliabilities ranged from .478 for "Relations with Others" to .717 for "Demeanor." Median reliabilities for other scales were .689 for "Job Knowledge," .643 for "Judgment," .670 for "Initiative," .667 for "Dependability," .714 for "Attitude," and .584 for "Communication."

Minority analyses were conducted on the data. Female ratees, black male ratees, and Spanish surnamed male ratees were compared to white male ratees. The female officers were all white. Tests were done, based on t-values, between minority and majority mean values to determine if the mean values for minority groups were significantly different from those obtained from a sample of white male officers. The t-test represents a very liberal test of the differences rather than a conservative test (i.e., the probability of a Type I error is most likely greater than the significance level reported)

due to the number of tests performed on the same data set. Black males and white females were rated significantly lower (at an alpha level less than .05) than the comparison group of white male officers on the dimension "Job Knowledge." Black males were also rated significantly lower (at an alpha level less than .05) on the dimensions "Dependability" and "Demeanor," and were rated significantly lower at an alpha level less than .01 on the dimension "Communication." White females were rated significantly higher (at an alpha level less than .05) on the dimension "Communication" than were the white male officers.

A principal components analysis with an oblique (Bi-Quartimin) rotation was conducted on the ratings. A scree test was used to determine the number of factors for rotation and the components analysis was conducted on the ratings from agencies free of extreme halo error. The following components seemed to be represented in the set of ratings: Component 1 seemed to represent "Work Attitude" or the quality of work; Component 2 comprised "Technical Adequacy"; and Component 3 seemed to represent "Interpersonal Relations."

The supervisory scales developed for this project were compared to performance appraisal scales currently in use in various police agencies with regard to the various rating errors of leniency, halo, and central tendency. Performance ratings on scales using several dimensions of job performance were obtained in eight cities. A total of thirteen cities

provided ratings on a single, overall scale of job performance. In both instances, the rating data in the comparative analyses were based on a common set of ratees.

Results of the comparative analyses showed that there was generally greater range in the median average ratings for departmental scales, but the median values for both sets of scales were quite similar. Average inter-dimension correlations for each performance dimension indicated generally comparable coefficients for both sets of ratings with regard to the median value. The supervisory rating scales for the project showed generally less central tendency error, as evidenced by the larger median standard deviations in all comparisons.

The supervisory rating scales demonstrated satisfactory psychometric adequacy. The mean values did not indicate undue leniency, and the standard deviations indicated that a large portion of the scale was actually being used in the rating process. The halo values indicated that there was a substantial degree of intercorrelation among the scales, but numerous agencies demonstrated low halo while other agencies demonstrated high halo. With the exception of the dimension "Relations with Others," the supervisory reliabilities were acceptable. It was difficult to determine whether the problem with this dimension is truly a reliability problem, or simply an indication of the problem of one agency in using the scales. The minority analyses of the supervisory ratings showed few differences between minority and majority officers. The factor

analysis of the supervisory scales identified three dimensions which adequately accounted for most of the variance in the supervisory ratings. The comparisons of the supervisory scales and departmental rating scales with regard to the rating errors of leniency, halo, and central tendency indicated that the supervisory rating instrument was generally comparable to the departmental scales on the characteristics of leniency and halo. The leniency error was not severe as the typical dimension mean was relatively near the midrange of the scales. The supervisory dimensional scales were positively correlated to a fairly strong degree, indicating perhaps a greater amount of halo error existed than would be desired. The supervisory scales were more resistant to central tendency error than were departmental scales in general.

In order to gather data regarding the construct validity of the supervisory scales, the relationships between the supervisory scales and a variety of other performance measures were obtained in several cities. These other measures included pre-employment civil service tests, biodata such as tenure and educational level, overall evaluation scales, and various non-rating performance data. The patterns of relationships varied from department to department, suggesting that situational characteristics may moderate the expected set of predictor-criterion and criterion-criterion relationships to be obtained in any one department. Despite the variability in findings, the general results indicated that the supervisory rating

scales appeared to be measuring various components of the job performance of police patrol officers.

APPENDIX D

PERFORMANCE DESCRIPTION SCALES

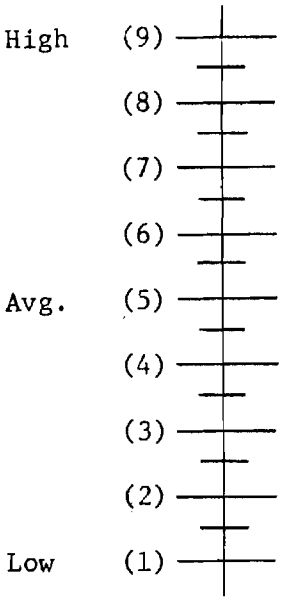
PERFORMANCE DESCRIPTION SCALES

City _____ Date _____

Supervisor _____ Officer _____

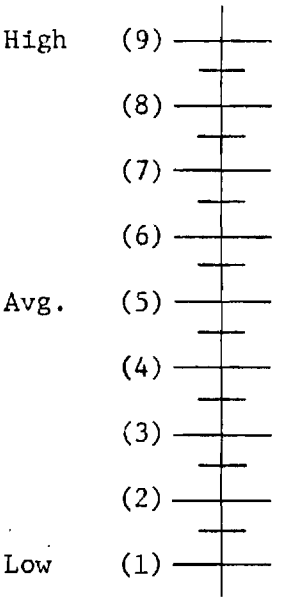
_____

Job Knowledge - Awareness of procedures, laws, and court rulings, and changes in them.



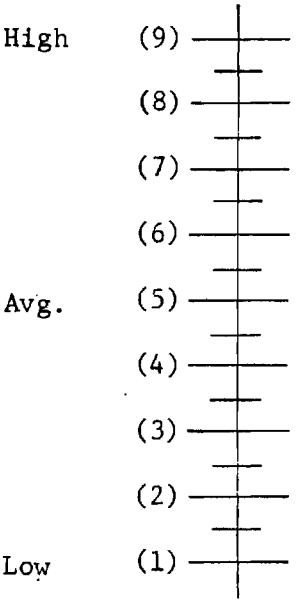
Comments:

Judgment - Observation and assessment of the situation and taking appropriate action.



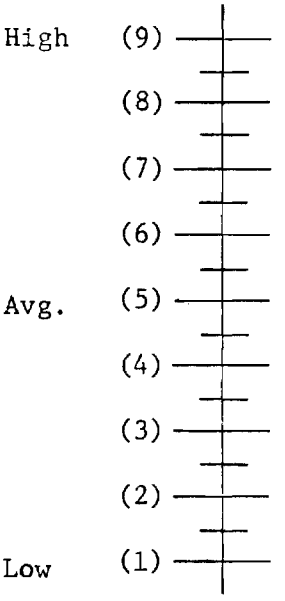
Comments:

Initiative - Individual personal performance conducted without either direct supervision or commands, including suggestions for improved departmental procedures.



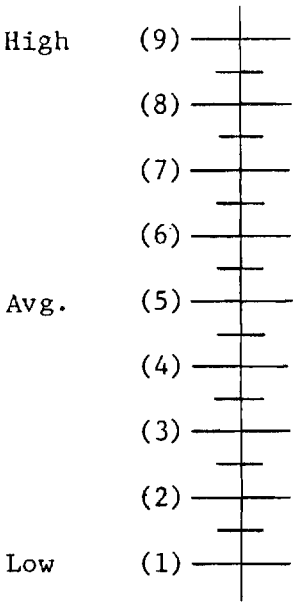
Comments:

Dependability - Predictable job behavior, including attendance, promptness, and reaction to boredom, stress, and criticism.



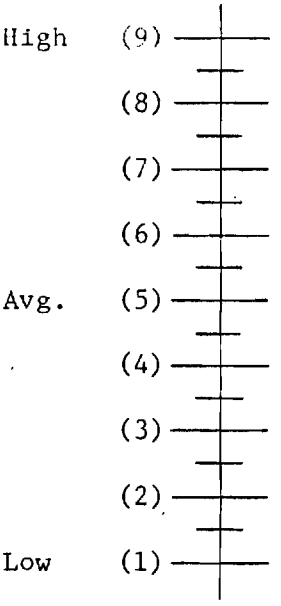
Comments:

Demeanor - Professional bearing as determined by overall neatness of uniform, personal grooming, and general physical condition.



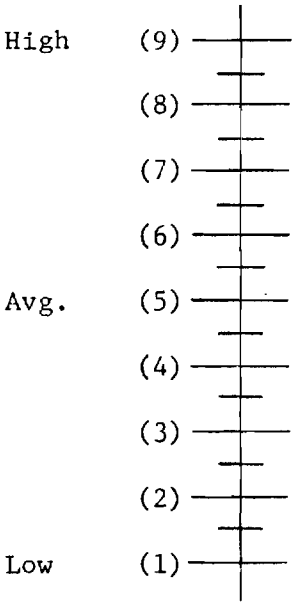
Comments:

Attitude - General orientation toward the law enforcement profession and the department.



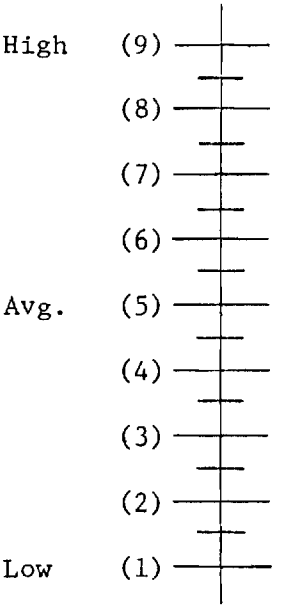
Comments:

Relations with others - Ability to deal with people he comes into contact with during the performance of his job, including the public, fellow officers, and supervisory personnel.



Comments:

Communication - Ability to make oneself understood and gather and transmit information, both in oral and written fashion.



Comments:

Summary Comments:

Do not write below this line.

JK	JU	IN	DP	DM	AT	RO	CO

APPENDIX E

INSTRUCTIONS AND EXAMPLES FOR PERFORMANCE DESCRIPTION SCALES (SUPERVISOR)

INSTRUCTIONS AND

EXAMPLES FOR

PERFORMANCE DESCRIPTION SCALES

(SUPERVISOR)

Instructions

The purpose of a rating scale is to provide an objective way of evaluating the "intangibles" of work performance. The rating scales which follow have been designed to help you rate the performance of your officers on eight areas which have been consistently identified as important aspects of patrol officer performance.

The rating process is simple. Consider the individual to be rated and give a rating on each of the eight (8) aspects of performance. You will notice that each of the eight aspects is defined for you. You will rate each officer by describing how well the officer typically performs the job aspect. When rating the officer, keep in mind the definition of the aspect being rated.

In other rating systems, it has been found that simply defining the work aspect for the rater does not provide enough information for accurate and reliable rating. Additional information is needed about the points along the rating scale. You will find this additional information starting on page 4 of this booklet.

Look at the fourth page of this booklet; you will see that a series of examples of job performance related to Job Knowledge are presented. These examples are placed in three broad categories: high, average, and low. These examples have been provided by first-line supervisors of patrol officers from agencies all around the country. The only examples provided are those which large numbers of supervisors agreed to put in a particular category (high, average, or low). We have found that these job performance examples help the supervisors make a more objective estimate of how well the officer has performed a particular job aspect during the rating period. We have also found that when supervisors use this kind of rating procedure, there is less confusion about the meaning of terms like "satisfactory," "outstanding," "needs attention," etc. which are usually found on rating scales.

You will notice that the rating scale itself is made up of a vertical line which has been numbered from 1 through 9. The purpose of the job performance examples starting on page 4 is to tell you exactly what level of performance is indicated by the various points along the scale. In other words, we are trying to

give you an idea of the behavior represented by a 7 or a 4 or a 2 on the scale.

The performance examples which have been provided by first-line supervisors of patrol officers may not adequately describe the performance of the individual patrol officer you are rating. If this is the case, construct your own performance example and write it in the "comments" section below the rating scale. There is one reaction which you might have to some job performance descriptions. Some raters find themselves disagreeing with the HIGH descriptions because they expect that level of performance from all of their officers. Consequently, they think that the descriptions should be considered AVERAGE. For example, if you require that all of your officers consistently preserve all potential evidence at the scene of a crime, this does not make the behavior average; IT MAKES YOUR REQUIRED LEVEL OF PERFORMANCE HIGH. In law enforcement, there are many areas in which nothing short of excellence or HIGH behavior is acceptable.

Steps in Rating an Officer

The actual rating procedure should be as follows:

1. Fill in your name and the name of the officer being rated on the first page of the rating form.

2. Turn to page 4 of this booklet labelled Job Knowledge; consider the definition of this aspect of performance; consider the examples of job performance which describe high, average, and low Job Knowedge, and decide on how much Job Knowledge the officer being rated has shown in the rating period.

3. Circle the point on the rating scale which best represents the amount of Job Knowledge demonstrated. This point can be any one of the hash marks along the scale (any one of the nine numbers or one of the halfway points between two numbers).

4. Turn to page 5 of the booklet, labelled Judgment, and rate the officer in a similar manner. Continue the process until the individual has been rated on each of the eight rating scales.

5. If you are rating more than one officer, repeat steps one through four until each officer has been rated.

Job Knowledge - Awareness of procedures, laws, and court rulings and changes in them.

The descriptions to the right are examples of behavior of individual patrol officers who are usually rated "High" on "Job Knowledge" by supervisors.

High

Always follows correct procedures for evidence preservation at the scene of a crime.

Is fully aware of recent court rulings, and conducts himself accordingly.

Searches a citizen's vehicle with probable cause, thereby discovering smuggled narcotics.

The descriptions to the right are examples of behavior of individual patrol officers who are usually rated "Average" on "Job Knowledge" by supervisors.

Average

Arrests a suspect at 11:00 P.M. on a warrant only after insuring that the warrant had been cleared for night service.

Distinguishes between civil matters and police matters.

Seldom has to ask others about points of law.

The descriptions to the right are examples of behavior of individual patrol officers who are usually rated "Low" on "Job Knowledge" by supervisors.

Low

Is consistently unaware of general orders and/or departmental policy.

Arrests a suspect for a misdemeanor not committed in his presence.

Misinforms the public on legal matters through lack of knowledge.

Judgment - Observation and assessment of the situation and taking appropriate action.

The descriptions to the right are examples of behavior of individual patrol officers who are usually rated "High" on "Judgment" by supervisors.

High

Calls for assistance and clears the area of bystanders before confronting a barricaded, heavily-armed suspect.

Notices potentially dangerous situations before anything actually occurs.

Radios in his position and discontinues a high-speed chase before entering areas of high vehicle and pedestrian traffic, such as school areas.

The descriptions to the right are examples of behavior of individual patrol officers who are usually rated "Average" on "Judgment" by supervisors.

Average

Issues warnings instead of tickets for traffic violations which occur at particularly confusing intersections for motorists.

Permits traffic violators to explain why they violated the law and then decides whether or not to issue a citation.

Does not leave a mother and daughter in the middle of a fight just because no law is being violated.

The descriptions to the right are examples of behavior of individual patrol officers who are usually rated "Low" on "Judgment" by supervisors.

Low

Enters a building with a broken door window instead of guarding the exits and calling for a backup unit.

Does nothing in response to a complaint about a woman cursing loudly in a restaurant.

Continues to write a traffic violation when he hears a report of a nearby robbery in progress.

Initiative - Individual personal performance conducted without either direct supervision or commands, including suggestions for improved departmental procedures.

The descriptions to the right are examples of behavior of individual patrol officers who are usually rated "High" on "Initiative" by supervisors.

High

Makes a special effort to find burglaries on his beat by carefully inspecting for signs of possible break-ins before the owners open their stores for business.

Comes to work early in order to check on the previous day's activities.

Seeks information about recent court rulings during his off-duty hours so that "good" arrests won't be lost by his actions.

The description to the right is an example of behavior of individual patrol officers who are usually rated "Average" on "Initiative" by supervisors.

Average

Fails to recognize and correct his own deficiencies without prompting from others.

The descriptions to the right are examples of behavior of individual patrol officers who are usually rated "Low" on "Initiative" by supervisors.

Low

Rarely checks the files for a suspect's friends and favorite hangouts.

Relies on his supervisor to make most of the important decisions for him.

Dependability - Predictable job behaviors, including job attendance, promptness, and reaction to boredom, stress, and criticism.

The descriptions to the right are examples of behavior of individual patrol officers who are usually rated "High" on "Dependability" by supervisors.

High

Remains cool under any circumstances.

Follows instructions.

Always gets to the station in time to check the daily log.

The descriptions to the righ
examples of behavior of ind
patrol officers who are us
rated "Average" on "Deper
by supervisors.

Average

inimum of sick days each

nough he has a
ticipated in the

The descrir
examples
patrol o
rated "
by super

Low

on receiving an emergency call.

e condition of his squad car as
use to avoid responding to a call.

cks up" in tense situations and
uts at other officers.

Demeanor - Professional bearing as determined by overall neatness of uniform, personal grooming, and general physical condition.

The descriptions to the right are examples of behavior of individual patrol officers who are usually rated "High" on "Demeanor" by supervisors.

High

Is meticulous about personal hygiene.

Works to keep himself in shape even though he's 45 years old.

Wears a clean, pressed uniform.

The descriptions to the right are examples of behavior of individual patrol officers who are usually rated "Average" on "Demeanor" by supervisors.

Average

Polishes boots and brass every day.

Cleans out his squad car at the end of a shift.

The descriptions to the right are examples of behavior of individual patrol officers who are usually rated "Low" on "Demeanor" by supervisors.

Low

Wears a uniform with holes in it.

Gets so fat that he can no longer do his job properly.

Reports for duty with his hair uncombed and an obvious hangover.

Attitude - General orientation toward the law enforcement profession and the department.

The descriptions to the right are examples of behavior of individual patrol officers who are usually rated "High" on "Attitude" by supervisors.

High

Considers law enforcement a career, not just a job.

Takes part in a study of police officers' opinions being conducted by a local college.

The description to the right is an example of behavior of individual patrol officers who are usually rated "Average" on "Attitude" by supervisors.

Average

Seldom gripes about departmental procedures.

The descriptions to the right are examples of behavior of individual patrol officers who are usually rated "Low" on "Attitude" by supervisors.

Low

Refuses training because he already is an expert.

"Goes out of his way" to defy departmental regulations.

Relations with others - Ability to deal with people he comes into contact with during the performance of his job, including the public, fellow officers, and supervisory personnel.

The descriptions to the right are examples of behavior of individual patrol officers who are usually rated "High" on "Relations with others" by supervisors.

High

Takes the time to carefully answer a Rookie's questions.

Maintains friendly relations with the civilians in his patrol area.

Establishes good relations with the youth in his patrol area by answering their questions and letting them look at his car and equipment.

The description to the right is an example of behavior of individual patrol officers who are usually rated "Average" on "Relations with others" by supervisors.

Average

Is considered "one of the boys" on his watch or shift.

The descriptions to the right are examples of behavior of individual patrol officers who are usually rated "Low" on "Relations with others" by supervisors.

Low

Always has fellow officers riled up by his actions and remarks.

Uses racially-toned language in front of minority group members.

Communication - Ability to make oneself understood and gather and transmit information, both in oral and written fashion.

The descriptions to the right are examples of behavior of individual patrol officers who are usually rated "High" on "Communication" by supervisors.

High

Turns in reports which are neat, accurate, and well written.

Carefully separates opinion from fact in his written and oral reports.

Uses correct grammar, spelling, and punctuation in written reports.

The description to the right is an example of behavior of individual patrol officers who are usually rated "Average" on "Communication" by supervisors.

Average

Never has to be asked to repeat himself over the radio.

The descriptions to the right are examples of behavior of individual patrol officers who are usually rated "Low" on "Communication" by supervisors.

Low

Includes far too much trivial, irrelevant information in his written reports and radio communications.

Uses "choppy," incomplete language in his written reports.

Talks so fast over the radio that he is unintelligible.

BIBLIOGRAPHY

Abbatiello, A. A study of police candidate selection. Paper presented at the 77th Annual Convention of the American Psychological Association, Washington, D. C., 1969.

Allen, G. M. The police manager's role in training. Police Chief, 1976, 43, 36-39.

Amick, D. J., and Crittenden, K.S. Analysis of variance and multivariate analysis of variance. In D. J. Amick and H. J. Walberg (Eds.), Introductory multivariate analysis. Berkeley, California: McCutchan Publishing Corporation, 1975.

Anastasi, A. Psychological testing. New York: Macmillan Publishing Company, 1976.

Anderson, H. E. Regression, discriminant analysis, and a standard notation for basic statistics. In R. B. Cattell (Ed.), Handbook of multivariate experimental psychology. Chicago, Illinois: Rand-McNally and Company, 1966.

Appelbaum, M. I. The MANOVA manual: Complete factorial design. Chapel Hill, North Carolina: L. L. Thurstone Psychometric Laboratory, University of North Carolina, 1974.

Appelbaum, M. I., and Cramer, E. M. Some problems in the nonorthogonal analysis of variance. Psychological Bulletin, 1974, 81, 335-343.

Atlanta Regional Commission. Test validation project report: Volume I. Atlanta, Georgia: Atlanta Regional Commission Governmental Service Department, October 1974.

Azen, S. P., Snibbe, H. M., and Montgomery, H. R. A longitudinal predictive study of success and performance of law enforcement officers. Journal of Applied Psychology, 1973, 57, 190-192.

Badalamente, R. V., George, C. E., Halterlein, P. J., Jackson, T. T., Moore, S. A., and Rio, R. Training police for their social role. Journal of Police Science and Administration, 1973, 1, 440-453.

Baehr, M. E., Furcon, J. E., and Froemel, E. C. Psychological assessment of patrolman qualifications in relation to field performance. Washington, D. C.: U.S. Government Printing Office, 1968.

Baehr, M. E., Saunders, D. R., Froemel, E. C., and Furcon, J. E. The prediction of performance for black and for white police patrolmen. Professional Psychology, 1971, 2, 46-57.

Balch, R. W. The police personality: Fact or fiction? Journal of Criminal Law, Criminology and Police Science, 1972, 63, 106-119.

Barrett, R. S., Taylor, E. K., Parker, J. W., and Martens, L. Rating scale content: I. Scale information and supervisory ratings. Personnel Psychology, 1958, 11, 333-346.

Baughman, E. E., and Dahlstrom, W. G. Negro and white children: A psychological study in the rural south. New York: Academic Press, 1968.

Beck, G. N. Municipal police performance rating. Journal of Criminal Law, Criminology and Police Science, 1961, 51, 567-574.

Bent, A. E. The politics of law enforcement. Lexington, Massachusetts: D. C. Heath & Company, 1974.

Bloch, P. B., and Anderson, D. Policewomen on patrol: Final report. Washington, D.C.: The Police Foundation, 1974.

Bock, R. D., and Haggard, E. A. The use of multivariate analysis of variance in behavioral research. In D. K. Witla (Ed.), Handbook of measurement and assessment in behavioral sciences. Reading, Massachusetts: Addison-Wesley, 1968.

Bodmer, W. F. Race and IQ: The genetic background. In M. Richards, K. Richardson, and D. Spears (Eds.), Race and intelligence: The fallacies behind the race-IQ controversy. Baltimore, Maryland: Penguin Books, Inc., 1972.

Bopp, W. J. Police personnel administration: The management of human resources. Boston, Massachusetts: Holbrook Press, 1974.

Borman, W. C. Effects of instructions to avoid halo error on reliability and validity of performance evaluation ratings. Journal of Applied Psychology, 1975, 60, 556-560.

Borman, W. C., and Dunnette, M. D. Behavior-based versus trait-oriented performance ratings: An empirical study. Journal of Applied Psychology, 1975, 60, 561-565.

Bozza, C. M. Motivations guiding policemen in the arrest process. Journal of Police Science and Administration, 1973, 1, 468-476.

Brodsky, S. L. Psychologists in the criminal justice system. Urbana, Illinois: University of Illinois Press, 1973.

Brown, F. G. Principles of educational and psychological testing. New York: Holt, Rinehart, & Winston, 1976.

Buros, O. K. (Ed.) The seventh mental measurements yearbook. Highland Park, New Jersey: The Gryphon Press, 1972.

Campbell, J., Dunnette, M., Arvey, R., and Hellervik, L. The development and evaluation of behaviorally based rating scales. Journal of Applied Psychology, 1973, 57, 15-22.

Cascio, W. F., and Real, L. J. Educational standards for police officer personnel. Police Chief, 1976, 43, 54-55.

Cohen, B., and Chaiken, J. M. Police background characteristics and performance. New York: The Rand Corporation, 1972.

Cramer, E. M. MANOVA: A computer program for univariate and multivariate analysis of variance. Chapel Hill, North Carolina: L. L. Thurstone Psychometric Laboratory, University of North Carolina, 1973.

Cronbach, L. J. Educational psychology. New York: Harcourt, Brace, and World, Inc., 1963.

Department of Labor. Validation of employment tests by contractors and sub-contractors subject to provisions of executive order 11246. Federal Register, 1968, 33, No. 186.

DuBois, P. H., and Watson, R. I. The selection of patrolmen. Journal of Applied Psychology, 1950, 34, 90-95.

Dunnette, M. D. Personnel selection and placement. Belmont, California: Brooks-Cole Publishing Company, 1966.

Eldefonso, E., Coffey, A., and Grace, R. C. Principles of law enforcement. New York: John Wiley and Sons, 1968.

Equal Employment Opportunity Commission. Guidelines on employee selection procedures. Federal Register, 1970, 35, No. 149.

Executive Committee of the Division of Industrial-Organizational Psychology of the Americal Psychological Association. Motion of the Executive Committee of Division 14 of the American Psychological Association for Leave to File Brief and for Oral Argument as Amicus Curiae in the Washington vs. Davis Case, 1975.

Fleek, T. A., and Newnam, T. J. The role of the police in modern society. Police, 1969, 13, 21-27.

Frost, T. M. Selection methods for police recruits. Journal of Criminal Law, Criminology and Police Science, 1955, 46, 135-145.

Fryer, D. H., Henry, E. R., and Sparks, C. P. General psychology. New York: Barnes and Noble, 1965.

Furcon, J. E. Some questions and answers about police officer selection testing. Chicago, Illinois: University of Chicago Industrial Relations Center, 1972.

Germann, A. C. Changing the police--The impossible dream? Journal of Criminal Law, Criminology and Police Science, 1971, 62, 416-421.

Guilford, J. P. Psychometric methods. New York: McGraw-Hill, 1954.

Guion, R. M. Personnel testing. New York: McGraw-Hill, 1965.

Hankey, R. O. Personality correlates in a role of authority: The police. Dissertation Abstracts, 1968, 29, 952.

Hanushek, E. A. Education and race: An analysis of the educational production process. Lexington, Massachusetts: D. C. Heath & Company, 1972.

Harris, R. J. A primer of multivariate statistics. New York: Academic Press, 1975.

Harris, R. J. Hand computation of Bonferroni critical values. Perceptual and Motor Skills, 1976, 42, 515-521.

Harris, R. N. The police academy: An inside view. New York: John Wiley & Sons, Inc., 1973.

Hess, L. R. Police entry tests and their predictability of score in police academy and subsequent job performance. Dissertation Abstracts, 1973, 33, 5552.

Hogan, R. Successful Oakland police: A personological profile. Unpublished manuscript, The Johns Hopkins University, 1972.

Iannone, N. F. Supervision of police personnel. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1975.

Institute of Industrial Relations. The current dilemma in employee selection: Legal and practical constraints on the use of tests and credentials in the employment process. Berkeley, Cal.: Institute of Industrial Relations, 1971.

Jensen, A. R. Social class, race, and genetics: Implications for education. In H. F. Clarizio, R. C. Craig, and W. A. Mehrens (Eds.), Contemporary issues in educational psychology. Boston, Mass.: Allyn & Bacon, 1970.

Kelly, R. M., and Farber, M. G. Identifying responsive inner-city policemen. Journal of Applied Psychology, 1974, 59, 259-263.

Kennedy, W. A. Child psychology. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1971.

Kirk, R. E. Experimental design: Procedures for the behavioral sciences. Belmont, California: Brooks-Cole, 1968.

Klineberg, O. Negro-white differences in intelligence test performance: A new look at an old problem. American Psychologist, 1963, 18, 198-203.

Landy, F. J., and Farr, J. L. Police performance appraisal: Final Report, Phase I. Washington, D.C.: National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, and U.S. Department of Justice, January 1973.

Landy, F. J., and Farr, J. L. Police performance appraisal: Final Report. Washington, D. C.: National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, and U.S. Department of Justice, September 1975.

Landy, F.J., Farr, J. L., Saal, F. E., and Freytag, W. R. Behaviorally anchored scales for rating the performance of police officers. Journal of Applied Psychology, 1976, 61, 750-758.

Landy, F. J., and Trumbo, D. A. Psychology of work behavior. Homewood, Illinois: The Dorsey Press, 1976.

Larzelere, R. E., and Mulaik, S. A. Single-sample tests for many correlations. Psychological Bulletin, 1977 (In press).

Latham, G. P., Wexley, K. N., and Pursell, E. D. Training managers to minimize rating errors in the observation of behavior. Journal of Applied Psychology, 1975, 60, 550-555.

Leonard, V. A. Police patrol organization. Springfield, Illinois: Charles C. Thomas, 1970.

Lindzey, G., Hall, C., and Thompson, R. F. Psychology. New York: Worth Publishers, Inc., 1975.

Lockwood, H. C. Critical problems in achieving equal employment opportunity. Personnel Psychology, 1966, 19, 3-10.

Mayfield, H. Equal employment opportunity: Should hiring standards be relaxed? Personnel, 1964, 41, 8-17.

McCormick, E. J., and Tiffin, J. Industrial psychology. Englewood Cliffs, New Jersey: Prentice-Hall, 1974.

McManus, G. P., Griffin, J. I., Wetteroth, W. J., Boland, M., and Hines, P. T. Police training and performance study. Washington, D.C.: U.S. Government Printing Office, 1970.

McNamara, J. H. Uncertainties in police work: The relevance of police recruits' backgrounds and training. In D. J. Bordua (Ed.), The police: Six sociological essays. New York: John Wiley and Sons, Inc., 1967.

Mills, R. B. New directions in police selection. Paper presented in a symposium, Profile of the U.S. Police Officer, at the 80th Annual Convention of the American Psychological Association, Honolulu, Hawaii, 1972.

Mills, R. B., McDevitt, R. J., and Tonkin, S. Situational tests in metropolitan police recruit selection. Journal of Criminal Law, Criminology and Police Science, 1966, 57, 99-106.

Milner, M. The illusion of equality: The effect of education on opportunity, inequality, and social conflict. San Francisco: Jossey-Bass Inc., 1972.

Milton, C. H., Abramowitz, A., Crites, L., Gates, M., Mintz, E., and Sandler, G. Women in policing: A manual. Washington, D.C.: The Police Foundation, 1974.

Morrison, D. F. Multivariate statistical methods. New York: McGraw-Hill, 1976.

Mullineaux, J. E. An evaluation of the predictors used to select patrolmen. Public Personnel Review, 1955, 16, 84-86.

National Advisory Commission on Criminal Justice Standards and Goals. Report on police. Washington, D.C.: U.S. Government Printing Office, 1973.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., and Bent, D. H. Statistical package for the social sciences. New York: McGraw-Hill, 1975.

Niederhoffer, A. Behind the shield: The police in urban society. New York: Doubleday & Company, Inc., 1967.

Otis, A. S., and Lennon, R. T. Otis-Lennon mental ability test: Manual for administration. New York: Harcourt, Brace & World, Inc., 1967.

Peters, D. L., and McCormick, E. J. Comparative reliability of numerically anchored versus job-task anchored rating scales. Journal of Applied Psychology, 1966, 50, 92-96.

Pollack, N. C. Use of written tests for police officer selection. Public Personnel Review, 1964, 25, 124-128.

President's Commission on Law Enforcement and Administration of Justice. Task force report: The police. Washington, D.C.: U.S. Government Printing Office, 1967.

Rao, C. R. Advanced statistical methods in biometric research. New York: John Wiley & Sons, 1952.

Raynor, J.O. Motivation and career striving. In J. W. Atkinson and J. O. Raynor (Eds.), Motivation and achievement. Washington, D. C.: V. H. Winston & Sons, 1974.

Richardson, J. F. Urban police in the United States. New York: National University Publications, 1974.

Robinson, D. D. Predicting police effectiveness from self reports of relative time spent in task performance. Personnel Psychology, 1970a, 23, 327-345.

Robinson, D. D. Characteristics of an effective police training program. Police, 1970b, 15, 42-46.

Ryan, J. IQ--The illusion of objectivity. In M. Richards, K. Richardson, and D. Spears (Eds.), Race and intelligence: The fallacies behind the race-IQ controversy. Baltimore, Maryland: Penguin Books, Inc., 1972.

Saunders, C. B. Upgrading the American police: Education and training for better law enforcement. Washington, D. C.: The Brookings Institution, 1970.

Sax, G. Principles of educational measurement and evaluation. Belmont, California: Wadsworth Publishing Co., 1974.

Schubert, E. J. Effective team management: State of the art. Police Chief, 1976, 43, 58-60.

Shuey, A. M. The testing of Negro intelligence. New York: Social Science Press, 1966.

Smith, D. H. Police officer selection: A critical literature review. Paper presented at the Annual Meeting of the Western Psychological Association, San Francisco, California, April 1971.

Smith, P. C., and Kendall, L. M. Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. Journal of Applied Psychology, 1963, 47, 149-155.

Spencer, G., and Nichols, R. A study of Chicago police recruits. Police Chief, 1971, 38, 50-55.

Spreen, J. F. Tomorrow's police--Protectors or pursuers? Police, 1971, 17, 21-25.

Tatsuoka, M. M. Multivariate analysis: Techniques for educational and psychological research. New York: John Wiley & Sons, Inc., 1971.

Taylor, C. R. Standards for selection, training and education of police personnel. Police, 1969, 14, 36-40.

Terman, L. M. A trial of mental and pedagogical tests in a civil service examination for policemen and firemen. Journal of Applied Psychology, 1917, 1, 17-29.

Thorndike, R. L., and Hagen, E. Measurement and evaluation in psychology and education. New York: John Wiley and Sons, Inc., 1969.

Timm, N. H. Multivariate analysis with applications in education and psychology. Monterey, California: Brooks-Cole, 1975.

Tyler, L. E. Tests and measurements. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1963.

Tyler, L. E. The psychology of human differences. New York: Appleton-Century-Crofts, 1965.

Tyler, L. E. Individual differences: Abilities and motivational directions. New York: Appleton-Century-Crofts, 1974.

U.S. Bureau of the Census. Statistical Abstract of the United States. Washington, D.C.: U.S. Government Printing Office, 1971.

Vinacke, W. E. Foundations of psychology. New York: American Book Company, 1968.

Watson, N. A. An application of social-psychological research to police work: Police-community relations. Dissertation Abstracts, 1967, 28, 791-792.

Weiss, D. J. Multivariate procedures. In M.D. Dunnette (Ed.), Handbook of industrial and organizational psychology. Chicago, Illinois: Rand McNally College Publishing Company, 1976.

Weldy, W. O. Women in policing: A positive step toward increased police enthusiasm. Police Chief, 1976, 43, 46-47.

Westley, W. A. Violence and the police: A sociological study of law, custom, and morality. Cambridge, Massachusetts: The MIT Press, 1970.

Weston, P. B. Police organization and management. Pacific Palisades, California: Goodyear Publishing Company, Inc., 1976.

Whisenand, P. M. Police supervision: Theory and practice. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1976.

Wilson, J. Q. Dilemmas of police administration. Public Administration Review, 1968, 28, 407-417.

Wilson, O. W. Police planning. Springfield, Illinois: Charles C. Thomas, 1962.

Zelen, M., and Severo, N. C. Probability functions. In M. Abramowitz and I. A. Stegun (Eds.), Handbook of mathematical functions. Washington, D.C.: National Bureau of Standards, 1964.